

Contents

Comparing portfolios built from intrinsic-value-based stock picks versus market-cap-weighted indices Jack James Vaduva, Dumitru Alexandru Bodislav	5
The impact of shifting elasticity of substitution on international firm location: global macroeconomic implications Wataru Johdo.....	21
Social security contributions and growth constraints in Greece: breaking the fiscal rigidity trap Panagiotis Asimakopoulos.....	31
Predicting customer churn using Machine Learning: models and core approaches Mohammad Sadegh Salem	45
Have we truly transitioned from industrial to service-based economies? An analysis of the economic sectoral distribution Nora Luz Quiroz Santiago	55
Monetary-fiscal policy regimes and economic growth in Algeria: a Markov-switching analysis (1990Q2–2025Q4) Moussa Ayachi, Abdeldjalil Chellig	71
Does foreign aid reach the poorest? A multidimensional poverty analysis for Pakistan Noel Alter, Arslan Khalid	87
Mathematical modeling of nearshoring and reshoring in the reconfiguration of trade flows Rares Nitu	103
Analyzing the impact of agricultural growth on inclusive growth in Côte d'Ivoire: key challenges and opportunities Konan Jules Kouadio, Koffi Célestin N'Goran, Wei Chang	117

Credibility erosion and exchange rate pass-through in Türkiye Umit Bulut, Okan Dadali, Azra Asel Bayrak	135
The aggregate banking-sector footprint of financial education mandates: cohort exposure, community bank margins, and an arithmetic of detection Dan-Mihai Dima	149
The relationship between exchange rates, trade and economic growth in the Philippine context Jason C. Patalinghug	173
Effect of inclusive business models on multidimensional poverty in Burkina Faso Balaram Roula, Dr.N. Kubendran	187
Capitalization of Romanian companies and macroeconomic variables: friends or foes? Bogdan-Octavian Cozmâncă, Leonard-Dan Uzum	201
The role of investment and trade in promoting economic growth in Ghana Adel Shakeeb Mohsen	215
Beyond generational decay: a psychosocial account of Generation Z's changing relationship with achievement, education and work Raluca Iuliana Georgescu, Dumitru Alexandru Bodisav	233
Financial inclusion and sustainable development in Algeria: an empirical analysis using the ARDL Model (2004-2022) Djamila Bekhti, Mehdi Bouchetara, Mohammed Redha Bouchikhi, Leila Ismahane Bakbak	245
The Impact of Fiscal Policy on Inflation in Romania Andreea Craciun	267
Comparing regional and global spillovers to the GCC equity markets Moosa Yousuf	281
Tax pressure and governance in a spatially interdependent world: international evidence and Morocco's positioning Mounir Atlassi, Mohamed Karim, Ilham Dkhissi	309
Management and performance of territorial police structures: a conceptual framework based on indicators Gabriel Nicolae Uță	331

Note: The authors are responsible for the content of their articles and for obtaining necessary permissions.

Mircea Dinu

Tel.: (+4) 031.432.96.02

Fax: (+4) 021.210.73.10

E-mail: comenzi@edecon.ro

Data base indexation:

EconLit

<http://www.aeaweb.org>

Research Papers in Economics (RePEc)

<http://www.ideas.repec.org>

<http://econpapers.repec.org>

Directory of Open Access Journals (DOAJ)

<http://www.doaj.org>

EBSCO Publishing

<http://www.ebscohost.com>

**International Consortium for the Advancement
of Academic Publication (ICAAP)**

<http://www.icaap.org>

Cabell's Directories

<http://www.cabells.com>

CNCIS B+

www.econometeoreticasiaplicata.ro; www.ectap.ro

Reception of texts: economia.ta@edeconomica.com

ISSN 1841-8678 (Print)

ISSN 1844-0029 (Online)

Comparing portfolios built from intrinsic-value-based stock picks versus market-cap-weighted indices

Jack James VADUVA

London School of Economics, United Kingdom
jackvaduva@gmail.com

Dumitru Alexandru BODISLAV

Bucharest University of Economic Studies, Romania
alex.bodislav@ase.ro

Abstract. *This study examines whether a portfolio selected on intrinsic value differs structurally from a capitalization-weighted portfolio drawn from the same universe. A random sample of 29 S&P 500 firms was valued with a free-cash-flow-to-the-firm discounted cash flow model; the 15 most undervalued names formed an equally weighted intrinsic-value portfolio, compared against a capitalization-weighted benchmark built from the same sample. The intrinsic-value portfolio showed lower valuation multiples, a higher dividend yield, a sector mix tilted toward cyclical and staple industries, and markedly lower representation among the largest firms. A three-month return comparison is reported for illustration only.*

Keywords: intrinsic value, discounted cash flow, market efficiency, capitalization weighting, portfolio construction.

JEL Classification: G11, G12, G14.

1. Introduction

The debate between active and passive investing remains one of the most persistent in finance. Passive strategies replicate a market index, typically weighted by capitalization, while active strategies attempt to identify mispriced securities. Evidence that most active managers fail to beat their benchmark once costs are included has strengthened the case for passive investing, yet value investors continue to argue that market prices do not always reflect a company's underlying economic worth.

This raises a narrower question: how much does a portfolio actually change when it is built from intrinsic-value estimates rather than market capitalization? Existing research approaches this mostly through returns, measuring whether value strategies, active management, or alternative weighting schemes beat their benchmarks over long samples, and remains caught in an unresolved debate over whether the patterns found reflect risk or mispricing. Less attention has gone to a simpler, structural question: holding the universe of available stocks fixed, how does an intrinsic-value selection rule reshape a portfolio compared with capitalization weighting, regardless of whether it wins on returns? This study addresses that question directly.

Intrinsic value is estimated for a random sample of 30 S&P 500 companies (reduced to 29 once a bank is excluded, since a free-cash-flow-to-the-firm model does not suit bank financials) using a discounted cash flow (DCF) approach, all valued at a single formation date. The 15 most undervalued firms are combined into an equally weighted intrinsic-value portfolio, and a capitalization-weighted benchmark is built from the same 29 names. Holding the universe constant isolates the effect of the selection and weighting rule itself, rather than differences in which stocks happen to be available.

Three hypotheses are tested: H1, that the intrinsic-value portfolio shows stronger value characteristics (lower valuation multiples and a higher dividend yield) than the benchmark; H2, that it differs in sector allocation; and H3, that it is less represented among the largest firms in the sample.

The contribution is descriptive rather than a claim about which approach performs better. A random sample reduces selection bias and allows differences between the portfolios to be attributed to the valuation and weighting method rather than to which stocks happen to be available. Because any difference in realized returns can always be attributed to the assumed model of risk rather than genuine mispricing, returns are reported only as illustration, not as evidence for or against outperformance. The remainder of the article reviews the relevant literature, sets out the valuation and comparison methodology, presents results across the three structural dimensions plus an illustrative return comparison, and concludes with the study's implications and limitations.

2. Literature Review

The starting point for this study is the efficient market hypothesis (EMH), most influentially stated by Fama (1970: pp. 383), who describes an efficient market as one in which prices fully reflect available information, sorted into weak, semi-strong, and strong forms

depending on which information set is assumed to be reflected. Fama found that the weak and semi-strong forms held up reasonably well, while the strong form did not, since insiders retain access to information prices do not yet reflect. This matters here because semi-strong efficiency is the central argument against what this study attempts: if public financial data is already priced in, an independent DCF valuation built from the same public data should not reliably find mispriced stocks. The intrinsic-value rule tested in this study assumes the opposite: that market price is not always the best estimate of value.

Efficiency, however, cannot be tested independently of a model of expected returns. For a researcher to judge whether prices incorporate all available information, they first need a model of what returns should look like in a normal, efficient market, and that is the model Fama (1970) builds his own analysis on. If a test then appears to show the market inefficient, the problem may just as easily lie with that assumed model of returns as with the market itself. This is the “joint hypothesis problem,” and it has direct consequences for this study: any difference in returns between the two portfolios built here could always be attributed to the chosen model of normal returns rather than to a genuine mispricing, which is a central reason the study is framed as a descriptive comparison rather than a test of efficiency. There are also theoretical and empirical reasons to doubt perfect efficiency. Grossman and Stiglitz (1980) argue that if prices already reflected all information, no one would be paid to gather it, so some inefficiency must persist to reward analysis, which is a promising starting point for a study built around independent valuation. Empirically, Fama and French (1992: pp. 427) show that firm size and the book-to-market ratio explain differences in average returns that market beta cannot, describing the relationship between beta and average return as essentially flat once size is taken into account, while Lakonishok, Shleifer and Vishny (1994: pp. 1541) argue such patterns instead reflect “the suboptimal behavior of the typical investor,” who systematically overpays for high-growth stocks and underpays for value stocks. The same evidence supports both readings. Fama and French treat it as a reward for risk that keeps the market efficient under a riskier model of returns, while Lakonishok, Shleifer and Vishny treat it as a mistake the efficient-market view cannot easily explain, and the debate remains unsettled, which is precisely why a comparison of price against independently estimated value remains relevant.

The study also sits within the active-versus-passive debate. Sharpe's (1991) “Arithmetic of Active Management” shows, by accounting logic rather than data, that because active and passive investors together make up the entire market, the average actively managed dollar must underperform the average passively managed dollar once costs are subtracted. This is an identity that holds in every period rather than a pattern that happens to appear in some samples and not others. Carhart (1997: pp. 57) supports this empirically, finding that apparent fund manager skill is largely explained by common return factors and cost differences rather than genuine stock-picking ability, so that persistence in fund performance rarely reflects anything more than exposure to the same systematic patterns. Sharpe also anticipates the obvious objection that some studies appear to show active managers beating the index, arguing that such results typically trace to measurement problems, such as “passive” benchmarks that are not truly passive, survivorship bias in the active sample, or an equal-weighted fund average that tilts toward small companies relative to the market, rather than to a flaw in the underlying logic. His arithmetic, however, applies

only to the average active investor; it does not, and cannot, rule out a specific selection rule beating the benchmark, which is the space this study's intrinsic-value approach occupies. The study does not attempt to overturn Sharpe's logic; it only asks whether a valuation-based selection rule produces a portfolio that differs systematically from the capitalization-weighted default his argument reinforces as the standard passive benchmark.

Where the EMH argues against active selection, the value-investing tradition, tracing to Graham and Dodd and described by Lakonishok, Shleifer and Vishny (1994: pp. 1541) as the founders of value investing, argues that a security's intrinsic value, derived from its fundamentals, can diverge from its market price, and that buying below intrinsic value is profitable. This idea gained empirical support when Fama and French (1992) identified the book-to-market ratio as a strong predictor of returns (the "value premium"), later formalized in their three-factor model (Fama and French, 1993). Interpretation of the value premium remains disputed: Fama and French read it as compensation for risk, while Lakonishok, Shleifer and Vishny (1994) read it as the correction of investors' predictable mistakes. This study, using a small descriptive sample, cannot settle that dispute; it can only show that the two portfolios differ, leaving the reason (risk, genuine mispricing, or the weighting rule itself) open.

Intrinsic value is estimated here using discounted cash flow (DCF) valuation, following the free-cash-flow-to-the-firm approach set out by Damodaran (2012), where expected free cash flows are discounted at the weighted average cost of capital (WACC) to an enterprise value and then converted to equity value per share. This is the most direct way of turning Graham and Dodd's qualitative notion of intrinsic value into a reproducible number. DCF carries weaknesses that matter for this study's design. Because cash flows are forecast explicitly for only a limited period, typically five to ten years, the terminal value typically dominates the total estimate (Damodaran, 2012), and because the constant-growth terminal-value formula divides by the gap between the discount rate and the assumed growth rate, that gap shrinking even slightly can move the whole valuation substantially, making the result highly sensitive to the assumed discount rate and long-run growth rate. Penman and Sougiannis (1998) find that DCF and dividend-discount models produce larger valuation errors than accrual-earnings models over the periods they test, attributing this to the terminal-value problem; Penman (2001: pp. 15) notes that free cash flow is a "perverse valuation attribute" because a firm investing to build value depresses its own near-term cash flow, so that a growing, value-creating company can show low or negative free cash flow for years, pushing an even larger share of its value into the fragile terminal estimate. Kaplan and Ruback (1995), by contrast, find DCF estimates come within roughly ten percent of observed transaction values on average when cash flows are forecast carefully against a real transaction price to anchor them, performing at least as well as valuations based on comparable companies. The literature's overall lesson is not that DCF should be rejected, since it remains the standard tool in practice and the one closest to the concept of intrinsic value, but that a single DCF figure should not be treated as exact. This motivates two design choices adopted below: every valuation is accompanied by a sensitivity analysis rather than reported as a fixed number, and realized performance is treated only as illustration, since the selection behind it is itself built on uncertain estimates.

The benchmark used for comparison is a capitalization-weighted index, in which each company's weight equals its share of total market value. This is the approach underlying the S&P 500 and most index funds, requiring little rebalancing since weights adjust automatically as prices move, and, under CAPM assumptions, the portfolio offering the best available mix of risk and return. Its popularity follows directly from that property, and it is also the passive default the average active manager cannot beat after costs, as Sharpe's arithmetic showed above. The most prominent critique comes from the fundamental-indexation literature: Arnott, Hsu and Moore (2005: pp. 83) argue that if prices contain errors, meaning they move away from fair value even temporarily, then weighting by market capitalization automatically puts more money into overvalued stocks and less into undervalued ones, so that capitalization weighting "assuredly overweights the overvalued stocks and underweights undervalued stocks." They propose weighting companies by fundamental measures of size instead, such as book value, earnings, dividends, sales, and employment, and report that indices built this way have historically outperformed their capitalization-weighted equivalents while holding the same stocks at different weights.

This claim needs to be treated with some care, for two reasons. First, it only holds if prices are wrong in a way an investor can actually use: under semi-strong efficiency, where price is an unbiased estimate of value and firm size alone says nothing about whether a stock is over- or undervalued, pricing errors are as likely to run one way as the other, and capitalization weighting carries no built-in penalty. Under that reading, fundamental indexation is not really a distinct passive strategy but an active value strategy presented as passive, which places it, like the intrinsic-value rule tested here, on the disputed side of the efficiency debate rather than settling it. Second, Arnott, Hsu and Moore are principals in a firm that sells fundamental-index products, and their outperformance figures come from a long historical back-test, so their commercial interest and the usual caveats around back-tested results are worth bearing in mind. The underlying idea nonetheless has a longer, less commercially entangled history: Haugen and Baker (1991) argued years earlier, without a comparable commercial stake, that capitalization-weighted portfolios are not the optimal risk-return tradeoff and that alternative weightings can do better. This study takes from the literature its core mechanism, that capitalization weighting concentrates capital in the largest names and, if prices are wrong, tilts toward the overvalued, without relying on the reported outperformance, which motivates two of its comparison dimensions: sector composition and representation among the largest firms.

Together, these strands form a single argument: the EMH provides the case against active selection; Sharpe's arithmetic sets the capitalization-weighted benchmark as the default the average manager cannot beat; the value-investing tradition provides the disputed counter-case that price and value can diverge systematically; DCF is the tool used to estimate that divergence while carrying its own fragility; and critiques of capitalization weighting suggest the passive default itself may be structurally biased toward large, overvalued names. Existing research addresses this question, whether independent valuation can differ from market price in a systematic and useful way, almost entirely through returns, using large samples, long histories, and performance measures, and remains caught in the unresolved risk-versus-mispricing debate. Less attention has gone to the simpler, structural question this study takes on: holding the universe of stocks fixed, how does an intrinsic-

value selection rule reshape a portfolio's valuation characteristics, sector composition, and representation among large firms, independent of whether it beats the market on returns? The contribution is therefore descriptive and exploratory: by fixing the universe through random sampling and prioritizing structural characteristics over performance, the study makes the mechanism by which intrinsic-value selection reshapes a portfolio visible and reproducible, without claiming to resolve the risk-versus-mispricing debate or to demonstrate a superior strategy.

3. Research Methodology

This study follows a mixed quantitative-descriptive design. It is quantitative in its underlying data, including cash-flow forecasts, discount rates, implied values, portfolio weights, valuation ratios, and returns, but descriptive in aim, measuring two portfolios built from a fixed universe and comparing them along four predetermined dimensions rather than testing statistical significance. With a single formation date and fewer than 30 firms, the sample is too small to support inferential testing of returns, and, following the joint hypothesis problem discussed above, testing efficiency directly is not the intention regardless. The design instead makes the selection mechanism itself visible: how an intrinsic-value rule and a capitalization-weighting rule, applied to the same stocks, produce structurally different portfolios.

The central research question is whether an intrinsic-value-based portfolio differs systematically from a capitalization-weighted benchmark drawn from the same fixed universe. Three hypotheses operationalize this: H1 predicts the intrinsic-value portfolio shows stronger value characteristics (lower P/E, P/B, and EV/EBITDA, higher dividend yield); H2 predicts a different sector allocation; H3 predicts lower representation among the sample's largest firms. A fourth, non-hypothesized comparison reports realized holding-period returns for illustration only.

The universe is the S&P 500. Thirty companies were drawn using Excel's random-number generator, each company having been numbered 1–500 in advance, removing any possibility of favourable stock-picking. One name, Citigroup, was excluded because the free-cash-flow-to-the-firm (FCFF) model does not suit banks, whose interest expense is part of operating activity rather than financing. This leaves 29 non-bank large-cap firms. (An insurance broker, Brown & Brown, was retained since its fee-based revenue fits the FCFF template.) All valuations use a single formation date of 27 February 2026; the two portfolios are then tracked over a three-month window from 2 March to 2 June 2026. All financial-statement, forecast, and market data, including market capitalization, shares outstanding, beta, the risk-free rate, and the equity risk premium, came from a single source, LSEG Workspace, to keep inputs consistent across firms.

Free cash flow to the firm was forecast explicitly for each company over a multi-year horizon, built up from consensus revenue-growth expectations, historical and projected operating margins, and normalized tax rates, adjusted for capital expenditure and working-capital needs; the weighted average cost of capital (WACC) used to discount those flows was built firm-by-firm from each company's beta, a common equity risk premium, and the

prevailing risk-free rate. A terminal value covering cash flows beyond the explicit horizon was then added, using the constant-growth (Gordon growth) formula:

$$TV = [FCF_n \times (1 + g)] \div (WACC - g)$$

The perpetual growth rate g is set case-by-case for each firm, ranging across the sample from roughly 1% (e.g. Kraft Heinz, reflecting its weak recent record as a mature consumer-staples business) to roughly 2.7% (e.g. Amazon, reflecting its scale and continued expansion), always kept below long-run nominal GDP growth of approximately 4% so that the valuation does not explode as g approaches WACC. Discounting the explicit cash flows and the terminal value back to the present and summing gives enterprise value:

$$\text{Enterprise Value} = \sum_n [FCF_t \div (1+WACC)^t] + [TV_n \div (1+WACC)^n]$$

Equity value follows by adding cash, short-term investments, and long-term investments, and subtracting short- and long-term debt:

$$\text{Equity Value} = EV + \text{Cash \& STI} + LTI - LT Debt - ST Debt$$

Dividing equity value by shares outstanding gives the implied share price, and comparing it with the market price on the formation date gives the upside/downside used to rank and select stocks:

$$\text{Upside/Downside} = (\text{DCF Implied Share Price} \div \text{Market Share Price}) - 1$$

Because a single DCF figure is fragile, as the literature review discussed, every valuation was stress-tested with a sensitivity grid varying WACC and the perpetual growth rate by ± 1.0 percentage point in 0.5-point steps around each firm's own estimate, and portfolio membership was checked at the extremes of that grid (Section 4.2).

Two portfolios were then built from the same 29 names, using deliberately opposite construction rules so that any structural difference between them can be traced to the rule itself rather than to the underlying stocks. The intrinsic-value portfolio holds the 15 most undervalued firms, the top 50% by upside, a cutoff fixed in advance rather than tied to a particular outcome, each equally weighted at $1/15$ ($\approx 6.67\%$). Equal weighting was chosen over weighting by the size of the upside because the upside figure itself is sensitive to model assumptions, and ranking a portfolio by a fragile number would concentrate the biggest bets in the valuations the study is least confident in. The capitalization-weighted benchmark applies the opposite logic to the same 29 names: it holds every firm regardless of valuation, weighted by market capitalization at the formation date, replicating standard index construction. Because it is drawn from only 29 names rather than a broad index, it can concentrate weight in a few large companies to a degree unlikely across the full S&P 500. This is a small-sample effect discussed among the limitations below and revisited when Amazon's weight is examined in Section 4.4. The S&P 500 itself is shown alongside as background context only, since its roughly 500-stock universe is not comparable in scale to the 29-stock sample, so any gap against it reflects universe size as much as selection method.

The two portfolios are compared across four dimensions: valuation characteristics (P/E, P/B, EV/EBITDA, dividend yield, reported as both median and weighted average), sector

composition (GICS classification, expressed as percentage of portfolio weight), representation among large-cap firms (the count of intrinsic-value holdings falling in the top half of the sample by market capitalization), and performance (three-month holding-period return, reported only as illustration given the short window and small sample).

The study carries several limitations that bound its conclusions. The sample of 29 firms from a single random draw and single formation date cannot be generalized to the broader market, and the sampling universe (current S&P 500 membership) introduces a degree of survivorship bias, though this matters little over a three-month window. Consensus forecasts and market parameters were collected on 14 May 2026, roughly ten weeks after the observation window began, which introduces look-ahead bias that affects the illustrative return comparison but not the structural valuation, sector, and representation findings, which rely on data available at formation. DCF valuations further depend on subjective forecasting assumptions and are more reliable for mature, stable-cash-flow firms than for cyclical or high-growth names; a large gap between market price and DCF estimate should not automatically be read as mispricing, since the model may not capture every risk. Reported returns are gross, unbalanced, and cover only three months, too short, and too easily dominated by single-stock noise, to support any claim about long-run investment performance.

4. Results and Analysis

This section applies the methodology above to the random sample. Section 4.1 presents the DCF valuation output and the resulting ranking; Section 4.2 checks the ranking's stability under a sensitivity analysis; Sections 4.3–4.5 compare the portfolios along the three hypothesized structural dimensions; Section 4.6 reports realized returns as illustration; Section 4.7 relates the findings to the literature.

4.1. Valuation results and the under/overvaluation ranking

Each of the 29 companies was valued with the FCFF DCF model described above; the implied price was compared with the closing price on the formation date (27 February 2026) to give an upside/downside figure, and the 15 names with the largest upside form the intrinsic-value portfolio. Table 1 reports the full ranking.

Table 1. *DCF valuation results and the under/overvaluation ranking*

Rank	Ticker	Company	GICS sector	Implied price	Price 27 Feb	Upside	Sel.
1	AMCR.N	Amcor PLC	Materials	\$145.54	\$48.43	200.5%	✓
2	KHC.OQ	Kraft Heinz Co	Consumer Staples	\$59.44	\$24.61	141.5%	✓
3	EXPE.OQ	Expedia Group Inc	Consumer Discretionary	\$449.18	\$215.69	108.3%	✓
4	VLO.N	Valero Energy Corp	Energy	\$394.82	\$204.64	92.9%	✓
5	KMB.OQ	Kimberly-Clark Corp	Consumer Staples	\$209.87	\$111.44	88.3%	✓
6	HAL.N	Halliburton Co	Energy	\$66.28	\$36.00	84.1%	✓
7	GEHC.OQ	GE Healthcare Technologies	Health Care	\$149.36	\$84.27	77.2%	✓
8	DAL.N	Delta Air Lines Inc	Industrials	\$110.67	\$65.70	68.4%	✓
9	BRO.N	Brown & Brown Inc	Financials	\$112.47	\$71.82	56.6%	✓

Rank	Ticker	Company	GICS sector	Implied price	Price 27 Feb	Upside	Sel.
10	EBAY.OQ	eBay Inc	Consumer Discretionary	\$122.10	\$90.86	34.4%	✓
11	OTIS.N	Otis Worldwide Corp	Industrials	\$120.84	\$92.56	30.6%	✓
12	MAS.N	Masco Corp	Industrials	\$92.21	\$71.62	28.8%	✓
13	JBL.N	Jabil Inc	Information Technology	\$307.49	\$264.99	16.0%	✓
14	PG.N	Procter & Gamble Co	Consumer Staples	\$192.05	\$167.20	14.9%	✓
15	DD.N	Dupont De Nemours Inc	Materials	\$55.73	\$50.04	11.4%	✓
16	KMI.N	Kinder Morgan Inc	Energy	\$36.90	\$33.27	10.9%	
17	WM.N	Waste Management Inc	Industrials	\$262.30	\$240.84	8.9%	
18	CARR.N	Carrier Global Corp	Industrials	\$70.06	\$64.40	8.8%	
19	ORCL.N	Oracle Corp	Information Technology	\$155.11	\$145.40	6.7%	
20	DOV.N	Dover Corp	Industrials	\$225.93	\$225.50	0.2%	
21	ALB.N	Albemarle Corp	Materials	\$178.55	\$178.67	-0.1%	
22	UNP.N	Union Pacific Corp	Industrials	\$234.26	\$264.98	-11.6%	
23	AKAM.OQ	Akamai Technologies Inc	Information Technology	\$84.12	\$98.39	-14.5%	
24	GRMN.N	Garmin Ltd	Consumer Discretionary	\$211.60	\$252.83	-16.3%	
25	VEEV.N	Veeva Systems Inc	Health Care	\$118.51	\$182.01	-34.9%	
26	AMZN.OQ	Amazon.com Inc	Consumer Discretionary	\$95.62	\$210.00	-54.5%	
27	PWR.N	Quanta Services Inc	Industrials	\$249.27	\$563.08	-55.7%	
28	ISRG.OQ	Intuitive Surgical Inc	Health Care	\$177.20	\$503.51	-64.8%	
29	COST.OQ	Costco Wholesale Corp	Consumer Staples	\$319.90	\$1010.79	-68.4%	
--	C.N	Citigroup Inc	Financials	n/a	n/a	n/a	

Note: calculations based on data from LSEG Workspace.

The spread of mispricing is wide. Amcor screens as the most undervalued name, with an implied price more than triple its market price, and Kraft Heinz, Expedia, Valero, Kimberly-Clark, and Halliburton all show upside above eighty percent; Costco, Intuitive Surgical, Quanta Services, and Amazon look heavily overvalued at the other end, trading well above the value the model implies. The undervalued half is dominated by cyclical and mature businesses: energy (Valero, Halliburton), materials (Amcor, Dupont), consumer staples (Kraft Heinz, Kimberly-Clark, Procter & Gamble), and a handful of industrials. The overvalued half holds large growth and quality names such as Amazon, Costco, Intuitive Surgical, and Veeva. The three technology names in the sample split across the cut, with Jabil inside it and Oracle and Akamai outside, so the pattern is closer to a price-versus-value rule that happens to capture many cyclical and staple companies than to a strict sector rule; this connection between the valuation screen and the cash-flow profile of these businesses resurfaces in Sections 4.3 and 4.7. The gap between selection and exclusion is small for a few borderline names. Dupont, the last stock included, shows an 11.4% upside against Kinder Morgan's 10.9% as the first excluded, which is exactly why the sensitivity check below matters: if the ranking flipped easily near that boundary, the selection would be an artifact of the model's assumptions rather than a real signal in the data.

4.2. Sensitivity analysis results

Each valuation was recomputed across a grid varying the terminal growth rate and discount rate (an adverse case lowering growth and raising the discount rate by 0.5pp each; a favourable case doing the reverse), and portfolio membership was re-checked at both extremes.

Table 2. *Portfolio membership under the sensitivity scenarios*

Scenario	Names held	Dropped out	Newly entering
Base case (as reported)	15 of 15	--	--
Adverse (g -0.5pp, WACC +0.5pp)	14 of 15	PG	DOV
Favourable (g +0.5pp, WACC -0.5pp)	13 of 15	DD, JBL	KMI, WM

Note: calculations based on data from LSEG Workspace.

The portfolio holds up well: only one name changes in the adverse case and two in the favourable case, both times among the borderline firms identified above. The deeply undervalued names (Amcor, Kraft Heinz) and the heavily overvalued names (Amazon, Costco) stay on their respective sides in every scenario. The selection sorts the sample into a clearly cheap group and a clearly expensive group, with only two or three borderline names able to swap places, which is sufficient stability for the structural comparison that follows.

4.3. Dimension 1: Valuation characteristics (H1)

Four characteristics were collected for every holding: P/E, P/B, EV/EBITDA, and dividend yield, each reported as both a median (robust to a single extreme name) and a weighted average. For the three multiples, a weighted harmonic mean was used to keep large outlier values from dominating the result; dividend yield, which sums cleanly across holdings, uses an ordinary weighted average.

Table 3. *Valuation characteristics of the two portfolios*

Characteristic	Bench. median	Bench. wtd. avg	IV median	IV wtd. avg	H1 met?
P/E	23.3	29.1	18.6	18.9	Yes (lower)
P/B	4.4	5.6	2.9	2.5	Yes (lower)
EV/EBITDA	14.9	18.1	11.9	10.8	Yes (lower)
Dividend yield	1.3%	0.7%*	2.0%	2.3%	Yes (higher)

Note: *weighted-average dividend yield counts non-payers as zero; Amazon alone is 52% of the benchmark and pays no dividend. Calculations based on data from LSEG Workspace.

The intrinsic-value portfolio is cheaper on all three multiples and offers a higher dividend yield, on both the median and the weighted-average basis, so H1 is supported on every measure, and the differences are not small: on the weighted figures the benchmark's EV/EBITDA is about seventy percent higher than the value portfolio's, and its price-to-book is more than double. The benchmark's weighted dividend yield of 0.7% reflects Amazon's 52% weight and zero payout; counting only its dividend-paying constituents lifts that figure to about 1.6%, but since payers are less than half the benchmark by weight, the lower, all-constituent figure is the honest one. Even against that adjusted figure, the intrinsic-value portfolio's yield remains higher, and most of its holdings pay a dividend. Kraft Heinz, which sits in the intrinsic-value portfolio with a normalized P/E above 100 due to depressed normalized earnings rather than a high price, is a useful robustness check:

on every other measure it is among the cheapest names in the sample, with one of the lowest P/B ratios and one of the highest dividend yields, and removing it from both portfolios shifts the intrinsic-value portfolio's harmonic P/E from 18.9 to 17.9 and its EV/EBITDA from 10.8 to 11.0, with no change of direction in any result and almost no movement in the benchmark, since Kraft Heinz is only about half a percent of its composition. The conclusion, in other words, does not depend on a single name. Cheaper multiples are consistent with genuine undervaluation, but equally consistent with the value premium being a reward for risk, and a sample this small cannot distinguish the two readings; the finding supported here is the descriptive one, that intrinsic-value selection with equal weighting produces a portfolio whose valuation characteristics differ from the benchmark in the predicted direction, for reasons discussed further in Section 4.7.

4.4. Dimension 2: Sector composition (H2)

Each holding was classified by GICS sector and weights summed within sector, using market-cap weights for the benchmark and equal weights for the intrinsic-value portfolio.

Table 4. *Sector composition of the two portfolios*

GICS sector	Benchmark	Intrinsic value	Difference (IV – bench.)
Consumer Discretionary	54.6%	13.3%	-41.3 pp
Consumer Staples	15.7%	20.0%	+4.3 pp
Energy	3.4%	13.3%	+9.9 pp
Financials	0.4%	6.7%	+6.3 pp
Health Care	3.7%	6.7%	+2.9 pp
Industrials	9.8%	20.0%	+10.2 pp
Information Technology	11.3%	6.7%	-4.6 pp
Materials	1.1%	13.3%	+12.3 pp

Note: calculations based on data from LSEG Workspace.

The two sector profiles differ sharply, supporting H2, though part of the contrast is mechanical: because the intrinsic-value portfolio is equally weighted, each sector's weight is simply its number of holdings multiplied by one-fifteenth, so Consumer Staples and Industrials each reach twenty percent purely because they hold three names apiece. The benchmark concentrates by construction and equal weighting flattens by construction, so some of the gap in Table 4 reflects that mechanical fact rather than the selection rule itself. A large share of the apparent Consumer Discretionary tilt compounds this with a single-stock effect: Amazon alone is 52% of the benchmark, and the sector's 54.6% weight is, read literally, less a sector bet than a bet on one company; removing Amazon collapses the benchmark's Consumer Discretionary weight below five percent, and Consumer Staples (Costco, Procter & Gamble) becomes its largest sector instead. Even so, a genuine tilt survives once Amazon is set aside: the intrinsic-value portfolio remains overweight the cyclical, asset-heavy, lower-multiple sectors (Energy, Materials, and Industrials) and underweight Information Technology, an underweight that widens further once Amazon is removed, since Oracle then makes up a large share of the benchmark's remaining technology weight and the intrinsic-value screen never selected it. This is coherent with Section 4.3: the portfolio is cheaper because it leans toward cyclical sectors the market holds lightly and away from the expensive growth names dominating the benchmark. A sector-level Herfindahl index of 0.35 for the benchmark against 0.15 for the intrinsic-value portfolio (falling to 0.22 for the benchmark without Amazon) puts the concentration into

one number: the benchmark is more than twice as concentrated at the sector level despite both portfolios holding the same eight sectors, and much of that concentration traces to a single holding, a result that connects directly to the fundamental-indexation critique discussed in Section 2 and carries into the representation dimension below.

4.5. Dimension 3: Representation among large-cap firms (H3)

The 29 stocks were ranked by market capitalization and the 15 intrinsic-value holdings located within that ranking; the headline measure is how many fall in the top half by size.

Table 5. Representation among large-capitalization firms

Representation measure	Names not selected	Intrinsic-value portfolio
Holdings in top half by market cap (top 15)	10 of 14	5 of 15
Holdings in bottom half (rank 16–29)	4 of 14	10 of 15
Three largest firms captured (AMZN, ORCL, COST)	3 of 3	0 of 3
Share of total sample market value	85.6%	14.4%
Median market capitalization	\$81.5 bn	\$27.6 bn

Note: calculations based on data from LSEG Workspace.

The result is clear and H3 is supported. Only five of the fifteen intrinsic-value holdings sit in the sample's larger half by market capitalization, and none of the three largest firms (Amazon, Oracle, Costco) are represented at all. Since the intrinsic-value portfolio is drawn from the same universe as the benchmark, every one of its names appears in the benchmark by construction, so the two portfolios necessarily overlap, and the real question is where in the size distribution that overlap sits. By weight the answer is stark. Although every intrinsic-value holding is also present in the benchmark, those names account for only 14.4% of the benchmark's market value, while the 14 firms the screen does not select account for 85.6%, with Amazon, Oracle, and Costco alone making up roughly 71%; the overlap between the two portfolios sits almost entirely in the benchmark's most lightly weighted tail. The median intrinsic-value holding is a \$27.6bn company against \$81.5bn for the names not selected, and this gap needs little interpretation: the screen captures none of the three largest firms and overlaps with only 14.4% of the benchmark by weight. The screen and the index draw from opposite ends of the same size distribution, which is the clearest evidence in the study that intrinsic-value selection does not reproduce capitalization weighting, and in this sample does close to the opposite.

4.6. Dimension 4: Performance (illustrative)

The final comparison reports holding-period returns over the observation window (2 March – 2 June 2026), with the S&P 500 as reference. Because the window is roughly twelve weeks and the intrinsic-value portfolio holds only fifteen stocks, results are dominated by noise; this section is reported for completeness, not as evidence for or against either construction method.

Table 6. Holding-period returns over the observation window

Portfolio	Weighting	Return (2 Mar – 2 Jun 2026)
Cap-weighted benchmark (29 names)	Market cap	+17.6%
S&P 500 (price index, reference)	Market cap	+10.5%
Intrinsic-value portfolio (15 names)	Equal	–0.5%
Memo: equal-weighted full universe (29)	Equal	+4.0%

Note: price returns, dividends not reinvested. Calculations based on data from LSEG Workspace.

The benchmark won comfortably, but the reason ties back to Sections 4.4 and 4.5 rather than to a general case for capitalization weighting: Amazon rose about 22% over the window and, at 52% of the portfolio, drove roughly two-thirds of the benchmark's entire gain on its own, with Oracle (+68%) accounting for much of the rest. The benchmark did well because this particular window happened to be a good one for the handful of large names that dominate it, the same concentration that widened the structural gaps in Sections 4.3 to 4.5, only this time it helped rather than hurt. The intrinsic-value portfolio's near-flat result is the mirror image, precisely because equal weighting means no single holding can carry the whole portfolio's return: its best names (Jabil +41%, Valero +26%) were offset by its worst (GE Healthcare -26%, Otis -24%), netting out close to zero, while an equally weighted portfolio of all 29 names in the sample returned about 4% over the same window, suggesting that the specific selection, and not simply the choice to equal-weight, held the intrinsic-value portfolio back here. Flipping to a window where Amazon fell would, on this logic, likely invert the ranking entirely. The S&P 500 is included only as background: it holds roughly 500 stocks against 29 in either portfolio here, so part of any gap against it reflects the difference in universe size rather than the selection rule, and says little about whether either method adds value. Over this one short period the benchmark outperformed, driven almost entirely by concentration in a single stock, and the study draws no conclusion from it about which construction method is superior, which is exactly why performance is framed throughout as illustration only.

4.7. Discussion: relating the findings to the literature

Together, the four dimensions tell one story: the intrinsic-value portfolio is cheaper on independent multiples, tilted toward cyclical and staple sectors and away from technology, and under-represented among the largest firms. These findings appear to share a common mechanism. The heaviest capitalization weights in this universe sit in a handful of large, highly priced growth names, above all Amazon, and the DCF screen avoids them, which mechanically produces a cheaper, differently weighted, smaller-cap portfolio. The performance result in Section 4.6 is the same mechanism viewed from another angle: the very concentration that widens the structural gaps happened, in this particular window, to favour the benchmark.

Set against the EMH (Section 2), the findings sit on the disputed side without settling it. Under semi-strong efficiency, an independent valuation built from public data should mostly agree with market prices, yet the intrinsic-value portfolio differs systematically and predictably from the benchmark drawn from the same stocks. This is consistent with Grossman and Stiglitz's (1980) argument that some inefficiency must persist to reward analysis, and with Graham and Dodd's premise that price and value can drift apart. But the joint hypothesis problem means the finding cannot on its own prove inefficiency, since any gap can be attributed to the assumed model of returns instead; the lower multiples documented are equally consistent with Fama and French's reading of the value premium as compensation for risk and with Lakonishok, Shleifer and Vishny's reading of it as the correction of investor mistakes, and a sample of 29 stocks over three months cannot distinguish between them.

The evidence speaks most clearly to the structure of the capitalization-weighted portfolio itself, connecting to Sharpe's arithmetic and to the fundamental-indexation critique. Sharpe's point, that the average active dollar must fall behind the average passive default, is not in question here; the study instead shows that a valuation rule builds a portfolio that differs meaningfully from that passive default. The benchmark's defining feature turned out to be concentration: one stock at roughly 52% of the portfolio and the three largest names above 70% of its value, the same structural tilt Arnott, Hsu and Moore describe when they argue capitalization weighting channels money toward the largest names and, if prices contain errors, toward the overvalued ones. The study adopts their central mechanism while remaining neutral on their performance claim, consistent with the decision to treat returns as illustration only.

The contribution is meant to be illustrative and exploratory rather than a generalizable discovery about markets. By fixing the universe through random sampling and prioritizing valuation characteristics, sector composition, and representation among large-cap firms ahead of performance, the study makes the effect of choosing intrinsic value over capitalization weighting visible on a scale that is easy to follow from start to finish, with the sensitivity analysis in Section 4.2 giving confidence that the effect reflects the selection rule itself rather than one fragile set of assumptions. It should be just as clear what the study cannot claim: it cannot settle the risk-versus-mispricing debate, cannot prove that value selection earns higher returns, and cannot generalize from 29 stocks drawn once from the S&P 500 to the entire market. Within those limits, the study can say that an intrinsic-value approach and capitalization weighting, applied to the same companies, produce structurally different portfolios in their valuation profile, sector exposure, and concentration among the largest firms, and that the difference traces consistently, across all three dimensions, to how each method treats the largest and most expensive names in the sample.

5. Conclusion

This study examined how a portfolio constructed from intrinsic-value estimates differs from a capitalization-weighted portfolio drawn from the same universe of stocks, focusing on structural differences rather than on resolving the long-standing debate over whether active management can consistently beat passive investing. A random sample of S&P 500 companies was valued with a free-cash-flow-to-the-firm DCF model; the fifteen companies with the largest discount between estimated intrinsic value and market price formed the intrinsic-value portfolio, compared against a capitalization-weighted benchmark from the same sample, so that any difference between the portfolios could be attributed to the selection and weighting rule rather than to differences in available stocks.

The results show the two approaches produce meaningfully different portfolios. The intrinsic-value portfolio displayed stronger value characteristics, lower valuation multiples and a higher dividend yield, supporting H1. It showed a markedly different sector composition, tilted toward Energy, Consumer Staples, and Industrials rather than the technology- and mega-cap-heavy benchmark, supporting H2. And it showed substantially lower representation among the sample's largest firms, with none of the three largest

companies selected and only 14.4% of the benchmark's market value overlapping with the intrinsic-value names, supporting H3.

Together these findings indicate that intrinsic-value selection and capitalization weighting are not simply alternative routes to the same portfolio: even applied to identical companies, the two approaches produce different valuation profiles, sector exposures, and concentration among large firms, driven by the construction rules rather than by the investment universe itself.

This study does not claim that one approach is superior. The short observation period, limited sample size, and inherent uncertainty of DCF valuation prevent broad conclusions about long-term performance, and the debates over market efficiency, value investing, and active management reviewed in Section 2 remain unresolved. The lower multiples and different sector tilt documented here are as consistent with a genuine value signal as with a reward for risk, and a sample of this size cannot arbitrate between the two. The contribution is meant to be descriptive and exploratory: by isolating the effects of selection and weighting methodology within a fixed universe, the study makes visible the structural consequences of choosing intrinsic value over market capitalization in portfolio construction, a necessary step before any claim about which approach delivers better long-term outcomes. Future work could extend the comparison across multiple random draws and formation dates to test whether the structural gap documented here (in valuation, sector mix, and representation among the largest firms) is a stable feature of intrinsic-value selection or an artifact of this particular sample, and could examine longer holding periods once look-ahead bias in the forecast inputs can be more fully controlled.

References

- Arnott, R.D., Hsu, J., and Moore, P. (2005). Fundamental Indexation, *Financial Analysts Journal*, 61(2), 83–99.
- Carhart, M.M. (1997). On Persistence in Mutual Fund Performance, *Journal of Finance*, 52(1), 57–82.
- Damodaran, A. (2012). *Investment Valuation: Tools and Techniques for Determining the Value of Any Asset*, 3rd ed. Hoboken, NJ: John Wiley & Sons.
- Fama, E.F. (1970). Efficient Capital Markets: A Review of Theory and Empirical Work, *Journal of Finance*, 25(2), 383–417.
- Fama, E.F. and French, K.R. (1992). The Cross-Section of Expected Stock Returns, *Journal of Finance*, 47(2), 427–465.
- Fama, E.F., and French, K.R. (1993). Common Risk Factors in the Returns on Stocks and Bonds, *Journal of Financial Economics*, 33(1), 3–56.
- Grossman, S.J., and Stiglitz, J.E. (1980). On the Impossibility of Informationally Efficient Markets, *American Economic Review*, 70(3), 393–408.
- Haugen, R.A., and Baker, N.L. (1991). The Efficient Market Inefficiency of Capitalization-Weighted Stock Portfolios, *Journal of Portfolio Management*, 17(3), 35–40.

- Kaplan, S.N., and Ruback, R.S. (1995). The Valuation of Cash Flow Forecasts: An Empirical Analysis, *Journal of Finance*, 50(4), 1059–1093.
- Lakonishok, J., Shleifer, A., and Vishny, R.W. (1994). Contrarian Investment, Extrapolation, and Risk, *Journal of Finance*, 49(5), 1541–1578.
- LSEG (2026). *Workspace* [online database]. London: London Stock Exchange Group. Available at: <<https://www.lseg.com/en/data-analytics/products/workspace>>
- Penman, S.H. (2001). On Comparing Cash Flow and Accrual Accounting Models for Use in Equity Valuation: A Response to Lundholm and O'Keefe, *Contemporary Accounting Research*, 18(4), 681–692.
- Penman, S.H., and Sougiannis, T. (1998). A Comparison of Dividend, Cash Flow, and Earnings Approaches to Equity Valuation, *Contemporary Accounting Research*, 15(3), 343–383.
- Sharpe, W.F. (1991). The Arithmetic of Active Management, *Financial Analysts Journal*, 47(1), 7–9.

The impact of shifting elasticity of substitution on international firm location: global macroeconomic implications

Wataru JOHDO

Tokyo Keizai University, Japan
johdo@tku.ac.jp

Abstract. *This study examines the impact of increased substitution elasticity between home and foreign tradable goods on firm location and macroeconomic variables. Using a two-country model, we show that deregulation in the tradable sector attracts producers to the home country and boosts local consumption and production, accompanied by currency appreciation. Although these consumption and production trends mirror those found in the non-tradable goods literature, our findings on firm relocation contrast sharply, revealing that tradable sector dynamics yield the opposite migratory effect.*

Keywords: elasticity of substitution, deregulation, tradable sector, production location, exchange rate.

JEL Classification: F31, F41, L51, L52.

1. Introduction

Since the 1990s, the globalization of economic activity—driven by deregulation aimed at removing barriers to trade and international capital flows—has accelerated the formation of global value chains (GVCs) in the manufacturing processes of firms headquartered in developed countries (Antràs, 2015; Tajoli and Felice, 2018; Kozul-Wright and Fortunato, 2019; Kano et al., 2020; Li et al., 2024). The expansion of GVCs has blurred the “nationality” of products and rendered the traditional, origin-based brand loyalty—the so-called Armington structure—obsolete. Domestically self-contained production systems are now the exception, and an international division of labor has become established, with developed countries handling Research and Development (R&D) and emerging and developing countries responsible for the processing and assembly of intermediate goods (Awate et al., 2015; Hertenstein and Williamson, 2018; Li, 2018). As a result, reverse imports—where products manufactured in emerging and developing countries flow back to developed countries as final consumption destinations—and cross-border imports and exports have become the standard forms of commodity circulation in the modern global economy. Furthermore, market liberalization and the convergence of product standards (Wang and Zhu, 2025), as well as the standardization of product specifications through compliance with global environmental regulations (Perkins and Neumayer, 2012), have reduced heterogeneity in the attributes of goods and increased the elasticity of substitution between domestic and foreign goods in consumers’ utility functions, thereby contributing to this trend. Furthermore, the rise of digital platforms has dramatically reduced international search costs through cross-border e-commerce, which is another important aspect worth noting (Lendle et al., 2016; Kim et al., 2017). In particular, the spread of online reputation feedback systems and return guarantees has reduced the uncertainty stemming from information asymmetry and is effectively eliminating the risk of adverse selection regarding low-quality foreign goods (Hui et al., 2016; Tadelis, 2016; Liu and Tang, 2018; Saeedi, 2019). These changes in the market environment have expanded consumers’ choice sets to a global scale, enabling them to compare prices and quality from a global perspective, and consequently leading to the establishment of consumer behavior that is sensitive to relative price differences. Furthermore, the advancement of transportation infrastructure, including maritime transport, has dramatically reduced trade costs—a factor that cannot be overlooked. This is because the reduction in trade costs, manifested as shorter lead times, has improved the effective geographical proximity between trading partners, further increased the substitutability of domestic and foreign goods, and served as a catalyst for market integration (Amador and Cabral, 2016).

The factors outlined above—including the deregulation of trade and international capital flows since 1990, the rise of digital platforms, the convergence of product standards, and the reduction in transportation costs—have dramatically increased consumers’ sensitivity to relative prices. This structural shift suggests that deregulation of trade goods within a single country amplifies the impact on trade volumes through fluctuations in exchange rates. In other words, deregulation in a global economy exposes domestic producers to

fierce international competition while providing consumers with welfare benefits in the form of access to cheaper goods. From a macroeconomic perspective, these phenomena can be formulated as an “increase in the elasticity of substitution between domestic and foreign goods,” thereby enabling an examination of their macroeconomic implications.

The purpose of this paper is to theoretically elucidate the impact on national macroeconomies of the “increased elasticity of substitution between domestic and foreign goods” against the backdrop of the structural changes in the global economy described above. For the analysis, we adopt the framework of the New Open Economy Macroeconomics (NOEM) which internalizes firms’ international location. By using this framework, we can clarify not only the traditional effects of increased elasticity of substitution on consumption and exchange rates but also the ripple effects it has on “firm relocation,” which progresses in parallel with economic globalization.

In terms of research related to this study, Cavelaars (2006) was one of the first researchers to analyze the effects of deregulation using the NOEM model. Cavelaars’s study elucidated the macroeconomic effects of deregulation in the non-tradable goods sector. The results showed that deregulation increases output in each country. Johdo (2019) used Cavelaars’s model to analyze the effects of deregulation in the non-tradable goods sector, taking into account the international location of firms. This analysis revealed that while deregulation increases output in the deregulated country, it also causes firms to relocate abroad due to currency appreciation. However, existing research has focused on the impact of deregulation as an increase in the elasticity of substitution between two goods in the non-tradable goods sector, and there has not yet been sufficient research on the impact of deregulation in the tradable goods sector corresponding to an increase in the elasticity of substitution between domestic and foreign goods.

In contrast to the two studies mentioned above, this study uses a NOEM framework that accounts for the cross-border location of producers to examine the macroeconomic effects of increased substitution elasticity between domestic and foreign goods in a given country—specifically, the effects of deregulation in the tradable goods sector. Our analysis shows that increased substitution elasticity between domestic and foreign goods leads to higher relative consumption and production in the deregulated country, as well as an appreciation of that country’s currency. Furthermore, producers relocate from the foreign country to the deregulated country. The resulting effects on production and consumption are identical to those demonstrated by Johdo (2019) for the non-tradable goods sector; however, the results regarding the international movement of producers in the tradable goods sector are the opposite of those in Johdo (2019). The results of this study suggest the following: If the objective of deregulation policies aimed at increasing the elasticity of substitution between domestic and foreign goods is to promote production and consumption, deregulation should be implemented regardless of whether the target industry produces tradable or non-tradable goods. However, if the government also intends to prevent domestic producers from relocating abroad—thereby mitigating concerns over

intellectual property infringement and supply chain risks, particularly during emergencies or disasters—deregulation should be prioritized for industries that produce tradable goods.

This paper is organized as follows. Section 2 presents the model. Section 3 presents the values of the key variables in a symmetric steady state. Section 4 describes the methods for the logarithmic-linear approximation of the model in both the short run and the long run. Section 5 uses the log-linear approximation methods described in Section 4 to clarify the macroeconomic effects of an increase in the elasticity of substitution between domestic and foreign goods—that is, the effects of deregulation in the tradable goods sector. Section 6 presents the conclusions.

2. Model

The model structure in this paper is based on the model by Johdo (2019). In this model, there are two countries: a domestic country and a foreign country. The asterisk symbol is employed to indicate foreign variables. Firms can move internationally, but their owners cannot. The intertemporal objective of agent i is to maximize the following utility function:

$$U_0(i) = \sum_{t=0}^{\infty} \beta^t \left(\log C_t(i) + \chi \log \left(\frac{M_t(i)}{P_t} \right) - \left(\frac{\kappa}{2} \right) (L_t^s(i))^2 \right), \quad (1)$$

where:

β ($0 < \beta < 1$) is a discount factor, $L_t^s(i)$ is the amount of labor effort, and the consumption index $C_t(i)$ is defined as $C_t(i) = \left(\int_0^1 C_t(i, j)^{(\sigma-1)/\sigma} dj \right)^{\sigma/(\sigma-1)}$, where $\sigma > 1$ and $C_t(i, j)$ is the good j 's consumption. In addition, $M_t(i)$ denotes nominal money balances and $P_t = \left(\int_0^1 P_t(j)^{1-\sigma} dj \right)^{1/(1-\sigma)}$. When ε_t represents the exchange rate, purchasing power parity is defined as $P_t = \varepsilon_t P_t^*$. In this model, producers in the range $[0, n_t]$ are located in the domestic country during period t , while the remaining firms in the range $(n_t, 1]$ are located in the foreign country. The representative agent's budget constraint can be expressed as follows:

$$P_t B_{t+1}(i) - P_t(1 + r_t)B_t(i) + M_t(i) - M_{t-1}(i) = \int_0^{n_t} \Pi_t(j) dj + \int_{n_t}^1 \varepsilon_t \Pi_t^*(j) dj + W_t(i)L_t^s(i) + P_t T_t(i) - P_t C_t(i), \quad (2)$$

where:

B_{t+1}^i denotes real bonds held by agent i , r_t denotes the real interest rate, W_t^i denotes the nominal wage rate, $\int_0^{n_t} \Pi_t(j) dj$ ($\int_{n_t}^1 \varepsilon_t \Pi_t^*(j) dj$) represents the total nominal profit flows of producers located in the domestic country (and abroad), and $T_t(i)$ denotes real lump-sum transfers.

In the domestic country, producer j , hires a continuum of differentiated labor inputs locally and produces a unique product according to $y_t(j) = \left(s^{-1/\phi} \int_0^s L_t^d(i, j)^{(\phi-1)/\phi} di \right)^{\phi/(\phi-1)}$, where $\phi > 1$.

The labor demand for agent i 's labor service, as determined by producer j , can be expressed as follows:

$$L_t^d(i, j) = s^{-1} (W_t(i)/W_t)^{-\phi} y_t(j), \text{ where } W_t \equiv \left(s^{-1} \int_0^s W_t(i)^{1-\phi} di \right)^{1/(1-\phi)}. \quad (3)$$

The product j 's market-clearing condition is as follows:

$$y_t(j) = \int_0^s C_t(i, j) di + \int_s^1 C_t^*(i, j) di = (P_t(j)/P_t)^{-\sigma} C_t^w, \quad (4)$$

where:

$C_t^w \equiv \left(\int_0^s C_t(i, j) di + \int_s^1 C_t^*(i, j) di \right)$. Similarly, we obtain $y_t^*(j) = (P_t^*(j)/P_t^*)^{-\sigma} C_t^w$. Agents seek to maximize (1) subject to (2).

The first-order conditions for this problem with respect to $B_{t+1}(i)$ and $M_t(i)$ can be written as $1/C_t(i) = \beta[(1 + r_{t+1})/C_{t+1}(i)]$ and $M_t(i)/P_t = \chi C_t(i)((1 + R_{t+1})/R_{t+1})$, respectively, where R_{t+1} is the nominal interest rate. In monopolistic competitive goods markets, domestic-located producer j faces the profit-maximization problem: $\max_{P_t(j)} \Pi_t(j) = (P_t(j) - W_t)y_t(j)$. By differentiating the profit with respect to $P_t(j)$, we obtain $P_t(j) = (\sigma/(\sigma - 1))W_t$.

Substituting (4) and the price mark-ups into $\Pi_t(j)/P_t$ and $\Pi_t^*(j)/P_t^*$, respectively, yields

$$\Pi_t(j)/P_t = (1/\theta)(P_t(j)/P_t)^{1-\theta} C_t^w, \quad \Pi_t^*(j)/P_t^* = (1/\theta)(P_t^*(j)/P_t^*)^{1-\theta} C_t^w. \quad (5)$$

In accordance with the Johdo (2019) formulation, the adjustment process for relocation of producers is presented as follows:

$$n_t - n_{t-1} = \gamma[\Pi_t(j)/P_t - \Pi_t^*(j)/P_t^*]. \quad (6)$$

The parameter γ ($0 \leq \gamma < \infty$) is the extent of producer mobility between two countries. In this study, the incorporation of monetary wage rigidities into the model is achieved through the implementation of single-period wage contracts. By substituting the equilibrium labor market conditions for the domestic country $L_t^s(i) = \int_0^{n_t} L_t^d(i, j) dj$ and equation (3) into (2), and maximizing (1) with respect to $W_t(i)$, the following first-order condition is obtained:

$$\phi(W_t(i)/P_t)^{-1} \left(\kappa(L_t^s(i))^2 \right) = (\phi - 1)(L_t^s(i)/C_t(i)). \quad (7)$$

3. Symmetric steady state

In the symmetric steady state, where $B_0 = 0$ and $s = s^* = 1/2$, the rate of real interest is given by $r_{ss} = (1 - \beta)/\beta \equiv \delta$. The allocation of producers is $n_{ss} = 1/2$. Hence, from (7), we obtain

$$L_{ss}^S = L_{ss}^{*S} = C_{ss} = C_{ss}^* = C_{ss}^W = y_{ss}(j) = y_{ss}^*(j) = \left(\frac{\phi-1}{\phi}\right)^{1/2} \left(\frac{\sigma-1}{\sigma}\right)^{1/2} \left(\frac{1}{\kappa}\right)^{1/2}. \quad (8)$$

Substituting C_{ss}^W from (8) into (5) yields the following levels of real profit:

$$\Pi_{ss}(j)/P_{ss} = \Pi_{ss}^*(j)/P_{ss}^* = \left(\frac{1}{\sigma}\right) \left(\frac{\phi-1}{\phi}\right)^{1/2} \left(\frac{\sigma-1}{\sigma}\right)^{1/2} \left(\frac{1}{\kappa}\right)^{1/2}. \quad (9)$$

4. Log-linear approximation

In order to examine the macroeconomic effects of a deregulation of the tradable goods sector in the domestic country, a log-linear approximation is solved around the zero-shock steady state. We use $\hat{Z}$ to denote the short-run percentage change of any variable Z that is consistent with the time span of monetary wage contracts. Furthermore, $\bar{Z}$ is employed to denote the long-run percentage change, a convention that aligns with the notion of flexible nominal wages.

5. Deregulation shocks of the tradable goods sector

We examine the effects of an unanticipated deregulation of the tradable goods sector on the exchange rate, producer relocation, and domestic consumption. In this model, the elasticity of substitution between two goods in the tradable goods industry (σ) is interpreted as a parameter that determines the degree of regulation of the tradable goods market in the domestic economy. Thus, a higher value of σ means that the tradable goods market is more deregulated. Therefore, a shock that increases σ can be considered as deregulation of the tradable goods market. Below, we examine the impact on the macroeconomy of an increase in σ ($\hat{\sigma} > 0$ and $\bar{\sigma} > 0$) in the first period. The log-linearization of equation (6) near the symmetric steady state yields the following short-run equation for the share of domestic producers:

$$\hat{n} = 2\gamma \left(\frac{\phi-1}{\phi}\right)^{1/2} \left(\frac{\sigma-1}{\sigma}\right)^{3/2} \left(\frac{1}{\kappa}\right)^{1/2} \left(\hat{\varepsilon} + \left(\frac{1}{\sigma-1}\right)\hat{\sigma}\right). \quad (10)$$

Equation (10) demonstrates how exchange rate depreciation affects producer behavior, specifically by shifting production locations toward domestic markets. This shift occurs because depreciation raises the relative cost of foreign production, thereby stimulating

domestic output through the expenditure-switching effect. This result is qualitatively the same as that in Johdo (2019). However, the second term in parentheses on the right-hand side of equation (10) does not exist in the corresponding equation in Johdo (2019), which focuses on deregulation of non-tradable goods industries. This term, $(1/(\sigma-1))\hat{\sigma}$, indicates that deregulation of tradable goods industries leads to producers moving from foreign countries to the domestic country through an increase in relative profits in the domestic country. Similarly, the long-run effect of an unexpected deregulation shock in the domestic country on the direction of international producer migration is as follows.

$$\bar{n} = 2\gamma \left(\frac{\phi-1}{\phi}\right)^{\frac{1}{2}} \left(\frac{\sigma-1}{\sigma}\right)^{\frac{3}{2}} \left(\frac{1}{\kappa}\right)^{\frac{1}{2}} \left[\bar{\varepsilon} - (\bar{W} - \bar{W}^*) + \left(\frac{1}{\sigma-1}\right) \bar{\sigma} \right]. \quad (11)$$

Similar to the result in (10), it is shown by (11) that producers are relocated from abroad to the domestic country by a depreciation of the domestic currency. Also, because nominal wages become more flexible in the long run, (11) is a decreasing function of the relative nominal wage differential in the domestic country. Similarly to the result in (10), this shows that an unexpected deregulation shock in the domestic country leads to a relocation of producers from abroad to the domestic country.

The macroeconomic effects of an unanticipated permanent deregulation of the tradable goods sector in the domestic country are now considered: $\hat{\sigma} = \bar{\sigma} > 0$. The following are the closed-form solutions for the three key variables:

$$\hat{C} - \hat{C}^* = \frac{2\delta\gamma_2\hat{\sigma} + \gamma_4\bar{\sigma}}{\delta\gamma_1 + \delta + \gamma_3} > 0, \quad (12)$$

$$\hat{\varepsilon} = - \left(\frac{2\delta\gamma_2\hat{\sigma} + \gamma_4\bar{\sigma}}{\delta\gamma_1 + \delta + \gamma_3} \right) < 0, \quad (13)$$

$$\hat{n} + \bar{n} = 2\gamma \left(\frac{\phi-1}{\phi}\right)^{1/2} \left(\frac{\sigma-1}{\sigma}\right)^{3/2} \left(\frac{1}{\kappa}\right)^{1/2} \left(\frac{2\delta\gamma_2\hat{\sigma} + \gamma_4\bar{\sigma}}{\delta\gamma_1 + \delta + \gamma_3} \right) \gamma_5 > 0, \quad (14)$$

where

$$\gamma_1 = \left(\frac{\sigma-1}{\sigma}\right)^2 \left[\frac{\sigma}{2} + 2\gamma \left(\frac{\phi-1}{\phi}\right)^{1/2} \left(\frac{\sigma-1}{\sigma}\right)^{1/2} \left(\frac{1}{\kappa}\right)^{1/2} \right] > 0,$$

$$\gamma_2 = \frac{1}{2} \left[1 + 4\gamma \left(\frac{1}{\sigma-1}\right) \left(\frac{\phi-1}{\phi}\right)^{1/2} \left(\frac{\sigma-1}{\sigma}\right)^{3/2} \left(\frac{1}{\kappa}\right)^{1/2} \right] > 0,$$

$$\gamma_3 = \left\{ 1 + \left(\frac{\sigma-1}{\sigma}\right) \left[\frac{\sigma-1+4\gamma \left(\frac{\phi-1}{\phi}\right)^{1/2} \left(\frac{\sigma-1}{\sigma}\right)^{3/2} \left(\frac{1}{\kappa}\right)^{1/2}}{\sigma+1+4\gamma \left(\frac{\phi-1}{\phi}\right)^{1/2} \left(\frac{\sigma-1}{\sigma}\right)^{3/2} \left(\frac{1}{\kappa}\right)^{1/2}} \right] \right\} > 0,$$

$$\gamma_4 = \frac{2}{\sigma} \left[\frac{\sigma + 4\gamma \left(\frac{\phi-1}{\phi}\right)^{1/2} \left(\frac{\sigma-1}{\sigma}\right)^{3/2} \left(\frac{1}{\kappa}\right)^{1/2}}{\sigma + 1 + 4\gamma \left(\frac{\phi-1}{\phi}\right)^{1/2} \left(\frac{\sigma-1}{\sigma}\right)^{3/2} \left(\frac{1}{\kappa}\right)^{1/2}} \right] > 0,$$

$$\gamma_5 = \left[\frac{\sigma + 4\gamma \left(\frac{\phi-1}{\phi}\right)^{1/2} \left(\frac{\sigma-1}{\sigma}\right)^{3/2} \left(\frac{1}{\kappa}\right)^{1/2}}{\sigma + 1 + 4\gamma \left(\frac{\phi-1}{\phi}\right)^{1/2} \left(\frac{\sigma-1}{\sigma}\right)^{3/2} \left(\frac{1}{\kappa}\right)^{1/2}} \right] > 0.$$

Equation (12) indicates that an unanticipated deregulation of the tradable goods sector in the domestic country results in a proportionate increase in the relative domestic consumption level. Regarding exchange rate changes, equation (13) shows that an unexpected deregulation of the tradable goods sector in the domestic country results in an appreciation of the exchange rate. The results of (14) demonstrate that when the domestic tradable goods industry is deregulated, producers will move from foreign countries to the domestic country. In particular, an intuitive explanation of equation (14) for international producer relocation is as follows: To begin, note that an unanticipated deregulation of the tradable goods sector in the domestic country has two opposing effects on the direction of the relocations of producers between countries. First, the unexpected deregulation of the tradable goods sector in the domestic country increases the global demand for that country's goods. This occurs through a decline in the price of those goods due to increased productivity. In turn, this increases production in the domestic country's tradable goods sector, ultimately leading to increased profits for producers in that sector. Hereafter, we shall call this the 'productivity effect'. Therefore, this productivity effect causes relocation of producers from the foreign country to the domestic country. Second, as shown in equation (13), an appreciation of the domestic currency shifts global demand from domestic products to foreign products. This shift in demand decreases the relative profits of firms located domestically, causing domestic producers to relocate to the foreign country. We refer to this as the 'exchange rate effect'. The direction of producer relocation between the two countries is determined by the relative strength of the productivity effect and the exchange rate effect. However, in our model, the productivity effect always outweighs the exchange rate effect, meaning that deregulation of tradable goods industries always leads to the movement of producers from foreign countries to the domestic market. This result is the exact opposite of that in Johdo (2019), whose analysis finds that deregulation of non-tradable goods industries results in the movement of producers from the domestic market to the foreign country.

6. Conclusion

This paper examined the impact of unexpected deregulation of a country's tradable goods sector on relative consumption, exchange rates, and producer mobility. The study compared the impact of unexpected deregulation of a country's tradable goods

sector with previous study predictions, particularly those concerning the international relocation of producers. Our results revealed that, unlike previous studies, when deregulation targets different industries, the impact on consumption and exchange rates remains consistent, while the direction of international producer mobility reverses.

References

- Amador, J., and Cabral, S. (2016). Global value chains: A survey of drivers and measures. *Journal of Economic Surveys*, 30(2), 278-301.
- Antràs, P. (2015). *Global production: Firms, contracts, and trade structure*. Princeton University Press.
- Awate, S., Larsen, M.M., and Mudambi, R. (2015). Accessing vs sourcing knowledge: A comparative study of R&D internationalization between emerging and advanced economy firms. *Journal of International Business Studies*, 46(1), 63-86.
- Cavelaars, P. (2006). Output and price effects of enhancing services sector competition in a large open economy. *European Economic Review*, 50(5), 1131-1149.
- Hertenstein, P., and Williamson, P.J. (2018). The role of suppliers in enabling differing innovation strategies of competing multinationals from emerging and advanced economies: German and Chinese automotive firms compared. *Technovation*, 70, 46-58.
- Hui, X., Saeedi, M., Shen, Z., and Sundaresan, N. (2016). Reputation and regulations: Evidence from eBay. *Management Science*, 62(12), 3604-3616.
- Johdo, W. (2019). Deregulation in non-tradable goods sector and relocation of firms in tradable goods sector. *Economics Bulletin*, 39(4), 2507-2516.
- Kano, L., Tsang, E.W., and Yeung, H.W.C. (2020). Global value chains: A review of the multi-disciplinary literature: Liena Kano et al. *Journal of International Business Studies*, 51(4), 577-622.
- Kim, T.Y., Dekker, R., and Heij, C. (2017). Cross-border electronic commerce: Distance effects and express delivery in European Union markets. *International Journal of Electronic Commerce*, 21(2), 184-218.
- Kozul-Wright, R., and Fortunato, P. (2019). Development unchained: Trade and industrialization in the era of international production. *Global Policy*, 10(1), 29-40.
- Lendle, A., Olarreaga, M., Schropp, S., and Vézina, P.L. (2016). There goes gravity: eBay and the death of distance. *The Economic Journal*, 126(591), 406-441.
- Li, L. (2018). China's manufacturing locus in 2025: With a comparison of "Made-in-China 2025" and "Industry 4.0". *Technological Forecasting and Social Change*, 135, 66-74.
- Li, T., Huang, X., and Du, D. (2024). Empirical research on the impact of international trade network of high-tech product on its global value chain. *The Journal of International Trade & Economic Development*, 33(6), 1040-1073.

- Liu, Y., and Tang, X. (2018). The effects of online trust-building mechanisms on trust and repurchase intentions: An empirical study on eBay. *Information Technology & People*, 31(3), 666-687.
- Perkins, R., and Neumayer, E. (2012) Does the 'California-Effect' operate across borders? Trading and investing-up in automobile emission standards. *Journal of European Public Policy*, 19(2), 217-237.
- Saeedi, M. (2019). Reputation and adverse selection: Theory and evidence from eBay. *The RAND Journal of Economics*, 50(4), 822-853.
- Tadelis, S. (2016). Reputation and feedback systems in online platform markets. *Annual Review of Economics*, 8(1), 321-340.
- Tajoli, L., and Felice, G. (2018). Global value chains participation and knowledge spillovers in developed and developing countries: An empirical investigation. *The European Journal of Development Research*, 30(3), 505-532.
- Wang, S., and Zhu, H. (2025). The adoption of international standards and export behavior of multi-product firms: A perspective based on differentiated competitive strategies. *Economic Analysis and Policy*, 86, 351-362.

Social security contributions and growth constraints in Greece: breaking the fiscal rigidity trap

Panagiotis ASIMAKOPOULOS

Athens University of Economics and Business, Greece
asimakop@aueb.gr, panasim86@yahoo.gr

Abstract. *This paper examines the macroeconomic role of labor taxation in Greece, focusing on the dynamic interaction between social security contributions (SSC) and real economic activity. Using a vector autoregressive framework, we document a pronounced asymmetry in the relationship between these variables. While increases in SSC exert a statistically and economically negative effect on GDP, feedback from GDP to SSC is negligible. The results indicate that SSC exhibits extreme persistence (0.95) and exerts a negative effect on GDP (-0.365), while GDP has negligible feedback on SSC. These findings suggest that SSC in Greece operates as a rigid, quasi-exogenous fiscal instrument that constrains growth. The results indicate that SSC behaves as a highly persistent and institutionally rigid policy variable, largely insulated from cyclical economic conditions. This asymmetry suggests that labor taxation operates as an exogenous constraint on growth rather than as an automatic stabilizer. Impulse response analysis and forecast error variance decomposition suggest that SSC shocks play a substantial role in driving output fluctuations. The findings contribute to the literature on taxation and macroeconomic performance by providing country-specific evidence on the structural role of labor tax rigidity and its implications for policy design in fiscally constrained economies.*

Keywords: social security contributions, fiscal rigidity, fiscal sustainability, structural reforms, Greece.

JEL Classification: H55, 047, E62.

1. Introduction

The relationship between taxation and economic performance has long occupied a central place in macroeconomic analysis⁽¹⁾. In economies characterized by fiscal imbalances and structural rigidities, the design of the tax system becomes particularly consequential for growth dynamics. Greece provides a compelling case in this regard. Following a prolonged period of economic adjustment, concerns persist regarding the role of labor taxation in constraining employment, competitiveness, and output. Among the components of the tax system, social security contributions (SSC) represent a critical but often underexamined element. As a major determinant of the labor tax wedge, SSC directly affects labor costs faced by firms and net wages received by workers. Despite its importance, empirical evidence on the dynamic macroeconomic effects of SSC remains limited, especially in the context of small open economies with institutional rigidities. This paper seeks to address this gap by analyzing the interaction between SSC and GDP using a vector autoregression (VAR) model. The empirical results reveal a striking asymmetry: SSC has negative effect on output, while output exerts virtually no influence on SSC. This pattern suggests that labor taxation in Greece is largely policy-driven and exhibits limited responsiveness to economic conditions. The contribution of this paper is twofold. First, it provides empirical evidence on the macroeconomic impact of labor taxation using a transparent and interpretable econometric framework. Second, it highlights the importance of institutional persistence in shaping fiscal dynamics, offering insights into why labor tax reform may be particularly effective as a growth-enhancing policy. The design of labor taxation occupies a central role in macroeconomic performance, particularly in economies characterized by structural rigidities and limited fiscal space. Greece represents a salient case in which labor tax burdens, especially those associated with social security contributions, have remained persistently high despite significant economic adjustment. The extent to which such taxation affects economic activity, however, remains an empirical question that requires careful identification and dynamic analysis. By combining reduced-form VAR analysis with a structural interpretation of policy shocks, the paper offers a coherent framework for understanding how institutional features of the tax system shape macroeconomic outcomes. The findings have direct implications for fiscal policy design and contribute to broader debates on the role of taxation in growth and stabilization.

2. Conceptual Framework

The theoretical mechanism underlying the analysis is straightforward but powerful. In standard macroeconomic models, output is produced using capital and labor, with firms choosing inputs to maximize profits. Labor taxation enters the model by driving a wedge between the cost of labor to firms and the net wage received by workers. Labor demand depends negatively on the effective wage, which includes social security contributions. An increase in SSC raises the effective cost of labor, reducing labor demand and, consequently, output. This mechanism is well established in the literature, but its empirical relevance

depends on institutional features such as wage rigidity, labor market flexibility, and the extent of informality. In contrast, the reverse channel—from output to SSC—is less clear. In principle, higher output could increase contributions through higher employment and wages. However, if SSC rates are set administratively and adjusted infrequently, this feedback mechanism may be weak or absent. The VAR framework provides a natural way to test these competing dynamics without imposing strong a priori restrictions.

3. Data, Econometric Methodology and Empirical Results

The estimated coefficients indicate that GDP depends positively on its own lag and negatively on lagged SSC. In contrast, SSC is almost entirely explained by its own lag, with minimal influence from GDP. The magnitude of the coefficients is economically meaningful. The persistence of GDP is moderate, suggesting that shocks dissipate over time. By contrast, SSC exhibits near-unit-root behavior, reflecting strong inertia. The negative coefficient of SSC in the GDP equation implies that increases in labor taxation lead to subsequent declines in output. This effect is consistent with the theoretical mechanism described earlier and provides empirical support for the hypothesis that labor taxation distorts economic activity. Equally important is the absence of a meaningful feedback effect from GDP to SSC. This indicates that labor taxation does not function as an automatic stabilizer in Greece. Instead, it appears to be determined largely by institutional and policy considerations. The model explains a substantial portion of the variation in SSC, reflecting its highly autoregressive nature. The explanatory power for GDP is more limited, which is not surprising given the omission of other macroeconomic variables such as investment, government spending, and external demand. Diagnostic checks suggest that the residuals are well behaved, and the system is dynamically stable, with all eigenvalues inside the unit circle. The empirical findings point to a structural asymmetry in the Greek economy. Labor taxation exerts a significant influence on output, but the reverse is not true. This asymmetry has important implications for both economic theory and policy. From a theoretical perspective, the results highlight the role of institutional rigidity in shaping macroeconomic dynamics. Standard models often assume that fiscal variables respond to economic conditions, but this assumption may not hold in practice. From a policy perspective, the findings suggest that reducing SSC could have a meaningful impact on economic activity. Given the high persistence of SSC, such reforms would need to be deliberate and sustained. The results support several policy recommendations. First, reducing social security contributions could stimulate economic activity by lowering labor costs and encouraging employment. This effect may be particularly strong in economies with high labor tax wedges. Second, introducing greater flexibility into the system of social security contributions could enhance macroeconomic stability. For example, linking contribution rates to economic conditions could improve the responsiveness of fiscal policy. Third, complementary reforms aimed at improving tax compliance and reducing informality could help offset potential revenue losses. While the

VAR framework provides valuable insights, it does not fully address issues of causal identification. Future research should explore structural VAR approaches, incorporate additional variables, and test for long-run relationships using cointegration techniques.

Table 1: *Coefficient Estimates*

Equation	Variable	Coefficient	Std. Error	t-statistic	Interpretation
GDP	GDP(-1)	0.5923	0.1163	5.0934	Strong positive persistence of GDP; previous period output explains current GDP (~59%). Statistically significant ($t > 2$).
GDP	SSC(-1)	-0.3650	0.2351	-1.5524	Negative effect of SSC on GDP. Not significant at 5%.
GDP	C	0.0398	0.0220	1.8126	Constant term, small positive drift in GDP over time.
SSC	GDP(-1)	0.0043	0.0194	0.2200	GDP has virtually no impact on SSC; t-statistic confirms insignificance.
SSC	SSC(-1)	0.9505	0.0392	24.2373	Very strong persistence of SSC; almost a unit root, reflects institutional rigidity in Greek social security policy. Highly significant.
SSC	C	0.0059	0.0037	1.6111	Constant term; small positive drift in SSC over time.

The estimated dynamics reveal a clear and economically meaningful structure in the interaction between output and labor taxation. Real GDP exhibits a moderate degree of persistence, with a coefficient of approximately 0.59 on its lagged value, indicating the presence of standard output inertia. This suggests that economic activity in Greece follows a path-dependent process in which past levels of production continue to influence current outcomes, although the effect is not strong enough to imply long-run rigidity or non-stationarity. In contrast, social security contributions display an exceptionally high degree of persistence, with a coefficient close to 0.95. This near-unit-root behavior points to a pronounced form of institutional rigidity in the determination of labor taxation. Changes in SSC appear to be highly inertial, evolving slowly over time and largely independent of short-term macroeconomic fluctuations. Such persistence is consistent with a policy environment in which contribution rates are determined through administrative and legislative processes rather than through automatic adjustment mechanisms. The interaction between the two variables is characterized by a significant asymmetry. Increases in SSC exert a negative short-run effect on GDP, providing empirical support for the theoretical prediction that higher labor taxation raises the cost of labor and dampens economic activity.

The magnitude of this effect is economically meaningful, suggesting that labor tax policy plays a nontrivial role in shaping output dynamics. At the same time, the reverse relationship is effectively absent. The estimated coefficient capturing the influence of GDP on SSC is extremely small, on the order of 0.004, and lacks economic significance. This indicates that changes in economic activity do not meaningfully feed back into the determination of labor taxation. Taken together, these results imply that social security contributions behave as a quasi-exogenous variable within the system. While they have a measurable and negative impact on output, they are not themselves influenced by fluctuations in GDP. This reinforces the interpretation of SSC as a policy-driven constraint on economic performance, rather than as an endogenous fiscal instrument that responds to the business cycle.

Table 2: Model Fit

Statistic	GDP	SSC	Interpretation
R-squared	0.4693	0.9399	GDP equation explains ~47% of variation; SSC equation explains ~94%, showing strong autoregressive structure.
Adjusted R-squared	0.4434	0.9369	Slightly lower, still robust.
F-statistic	18.1279	320.4628	Both equations are jointly significant; SSC equation extremely well explained.
S.E. equation	0.0263	0.00438	Standard error of regression; lower for SSC due to persistence.

The overall fit of the model reveals an important asymmetry in explanatory power across the two equations. Social security contributions are explained to a very high degree within the VAR framework, which reflects their strong autoregressive structure. The high goodness of fit is consistent with the earlier finding of near-unit-root persistence, indicating that the evolution of SSC is largely driven by its own past values. This pattern reinforces the view that labor taxation in Greece is governed by institutional inertia, with limited sensitivity to other macroeconomic variables included in the system. By contrast, the explanatory power of the GDP equation is more moderate. While past values of GDP and SSC do account for a meaningful portion of output fluctuations, a substantial share of variation remains unexplained within the bivariate framework. This is not unexpected, as economic activity is influenced by a broader set of macroeconomic forces that are not explicitly captured in the model. Factors such as private investment dynamics, government expenditure, and inflows from European Union programs are likely to play an important role in shaping output. The relatively lower fit of the GDP equation therefore reflects the inherent complexity of the growth process.

Table 3: Information Criteria

Criterion	GDP	SSC	Interpretation
AIC	-4.3752	-7.9570	Lower AIC indicates better fit; SSC equation very well fitted.
SC	-4.2535	-7.8354	Schwarz criterion consistent with AIC.

The estimation results indicate that the model provides an excellent fit for social security contributions and a moderate fit for GDP. The high explanatory power of the SSC equation reflects the strong persistence inherent in the series, while the somewhat lower fit of the GDP equation suggests that output is influenced by additional macroeconomic factors not captured within the bivariate framework. Examination of the residual covariance matrix reveals a determinant of 1.27×10^{-8} , a small value that indicates the residuals are well-behaved and that the system does not exhibit signs of instability. Moreover, the estimated VAR (1) coefficients are all less than one in modulus, implying that there are no explosive roots and that the dynamic system is stable. This diagnostic evidence reinforces the interpretation of the structural dynamics. Shocks to SSC dominate the behavior of the system due to the variable's extreme persistence, while GDP, though moderately influenced by its own history and by SSC, responds more actively to other disturbances. Despite the higher variability in GDP, the negative impact of SSC on output remains economically significant. The combined evidence underscores the asymmetric role of social security contributions within the Greek macroeconomy: SSC operates as a quasi-exogenous, highly persistent force that constrains growth, while GDP reacts to these shocks

without exerting meaningful feedback on the contribution system. This dynamic asymmetry highlights both the importance of SSC in driving macroeconomic outcomes and the structural rigidity that limits the responsiveness of labor taxation policy. The empirical results for Greece provide a clear economic interpretation of the role played by social security contributions in shaping macroeconomic outcomes. The extremely high persistence of SSC, with an estimated coefficient of 0.95, confirms the presence of strong policy rigidity. In practice, this means that contributions do not automatically adjust in response to changes in economic performance, rendering SSC largely insensitive to the business cycle. This rigidity reflects the institutional and political constraints embedded in the Greek fiscal system and highlights the quasi-fixed nature of labor taxation in the short run. The negative impact of SSC on GDP, estimated at approximately -0.365 , underscores the contractionary effects of high labor taxation. Elevated contributions increase the effective cost of labor, reduce competitiveness, and slow output growth. This effect is particularly salient in Greece, where a substantial labor tax wedge coincides with a sizeable informal sector, amplifying distortions and reducing incentives for formal employment. Firms face higher costs, small and medium-sized enterprises in particular are constrained, and overall economic expansion is dampened by the structural burden of contributions. At the same time, feedback from GDP to SSC is negligible, with an estimated coefficient near zero. This indicates that social security contributions behave as a quasi-exogenous policy variable; economic growth does not automatically alleviate contribution burdens, nor does a downturn trigger automatic reductions. The lack of countercyclical adjustment highlights a major limitation in the design of the Greek social security system, depriving the economy of an important stabilizing mechanism. The policy implications of these findings are substantial. Gradual reduction of SSC rates or a reform toward more flexible, countercyclical contributions could enhance growth by lowering labor costs and stimulating employment. Administrative reforms, such as the adoption of a Balanced Scorecard approach within the tax authority, could help offset potential revenue losses while improving efficiency. At the same time, targeted efforts to formalize the labor market would expand the tax base and reduce incentives for informality, creating a more sustainable and growth-friendly fiscal framework. From an academic and policy perspective, these results confirm the broader literature indicating that labor taxation can constrain output. The contribution of the Greek case lies in its demonstration of asymmetric VAR dynamics, where the rigidity of policy dominates economic feedback, producing persistent negative effects on output. By highlighting this structural asymmetry, the analysis provides both theoretical confirmation and practical guidance for evidence-based reform.

4. Conclusion

This paper has examined the dynamic relationship between social security contributions and economic activity in Greece. The results reveal a clear asymmetry in which labor taxation exerts a significant influence on output, while remaining largely unaffected by it.

These findings suggest that SSC functions as a persistent and policy-driven constraint on growth. Addressing this constraint through structural reform may therefore represent a key lever for improving macroeconomic performance. Future research should build on these results by incorporating stronger identification strategies and exploring the interaction between labor taxation and other dimensions of fiscal policy.

The Greek case provides a particularly relevant setting for examining the macroeconomic effects of labor taxation, given its historically high labor tax wedge, its strong reliance on social security contributions for pension financing, and the presence of persistent structural rigidities in tax policy. These features have been further shaped by the legacy of the sovereign debt crisis, which imposed significant fiscal constraints and led to a series of policy adjustments. Against this background, the central question arises as to whether social security contributions function as a structural barrier to economic growth in Greece. The institutional configuration of SSC helps to frame this question more precisely. In Greece, social security contributions are characterized by relatively high rates for both employers and employees, reflecting their central role in sustaining the pension system.

At the same time, the system lacks automatic adjustment mechanisms that would allow contributions to respond flexibly to changing macroeconomic conditions. During the crisis period from 2010 to 2018, policy reforms were primarily oriented toward securing fiscal sustainability. This objective was pursued through increases in contribution rates, as well as efforts to broaden the contribution base. While these measures were successful in stabilizing public finances, they may also have amplified distortions in the labor market by increasing the cost of formal employment and weakening incentives for job creation. The broader significance of the findings lies in their ability to move beyond a purely descriptive exercise. The results provide a clear diagnosis, identifying the persistence and growth-dampening effects of SSC as a central constraint within the Greek economy.

They also uncover the underlying mechanism, namely the asymmetric dynamic relationship between taxation and output, which reinforces this constraint over time. Most importantly, they point toward a solution that integrates structural tax reform with greater institutional flexibility and improved policy responsiveness. In conclusion, the evidence indicates that social security contributions in Greece are highly persistent, exert a negative influence on economic activity, and remain largely unresponsive to changes in macroeconomic conditions. This combination creates a self-reinforcing dynamic in which taxation constrains growth without adapting to it. Breaking this cycle requires a comprehensive approach that includes reducing the burden of labor taxation, introducing countercyclical features into the system, and strengthening administrative capacity. Such reforms have the potential not only to alleviate existing distortions but also to support a more dynamic and resilient growth trajectory.

Econometric Output

Vector Autoregression Estimates		
Date: 11/09/22 Time: 18:37		
Sample (adjusted): 1975 2018		
Included observations: 44 after adjustments		
Standard errors in () & t-statistics in []		
	GDP	SSC
GDP(-1)	0.592290 (0.11628) [5.09363]	0.004257 (0.01940) [0.21948]
SSC(-1)	-0.364955 (0.23509) [-1.55241]	0.950456 (0.03921) [24.2373]
C	0.039819 (0.02197) [1.81264]	0.005903 (0.00366) [1.61105]
R-squared	0.469295	0.939876
Adj. R-squared	0.443407	0.936943
Sum sq. resids	0.028291	0.000787
S.E. equation	0.026268	0.004382
F-statistic	18.12786	320.4628
Log likelihood	99.25384	178.0541
Akaike AIC	-4.375175	-7.957006
Schwarz SC	-4.253525	-7.835357
Mean dependent	0.015836	0.091030
S.D. dependent	0.035210	0.017449
Determinant resid covariance (dof adj.)		1.27E-08
Determinant resid covariance		1.10E-08
Log likelihood		278.2378
Akaike information criterion		-12.37445
Schwarz criterion		-12.13115
Number of coefficients		6

Roots of Characteristic Polynomial	
Endogenous variables: GDP SSC	
Exogenous variables: C	
Lag specification: 1 1	
Date: 11/09/22 Time: 18:39	
Root	Modulus
0.946065	0.946065
0.596681	0.596681
No root lies outside the unit circle. VAR satisfies the stability condition.	

VAR Granger Causality/Block Exogeneity Wald Tests			
Date: 11/09/22 Time: 18:40			
Sample: 1974 2018			
Included observations: 44			
Dependent variable: GDP			
Excluded	Chi-sq	df	Prob.
SSC	2.409974	1	0.1206
All	2.409974	1	0.1206
Dependent variable: SSC			
Excluded	Chi-sq	df	Prob.
GDP	0.048170	1	0.8263
All	0.048170	1	0.8263

VAR Lag Exclusion Wald Tests

Date: 11/09/22 Time: 18:40

Sample (adjusted): 1975 2018

Included observations: 44 after adjustments

Chi-squared test statistics for lag exclusion:
Numbers in [] are p-values

	GDP	SSC	Joint
Lag 1	36.25573 [0.0000]	640.9255 [0.0000]	672.3889 [0.0000]
df	2	2	4

VAR Lag Order Selection Criteria

Endogenous variables: GDP SSC

Exogenous variables: C

Date: 11/09/22 Time: 18:41

Sample: 1974 2018

Included observations: 41

Lag	LoqL	LR	FPE	AIC	SC	HQ
0	194.2848	NA	2.89e-07	-9.379748	-9.296159	-9.349310
1	261.0416	123.7443*	1.36e-08*	-12.44106*	-12.19029*	-12.34974*
2	261.7055	1.165731	1.60e-08	-12.27832	-11.86037	-12.12612
3	263.9437	3.712211	1.75e-08	-12.19238	-11.60725	-11.97931
4	267.1194	4.957111	1.84e-08	-12.15216	-11.39986	-11.87822

* indicates lag order selected by the criterion

LR: sequential modified LR test statistic (each test at 5% level)

FPE: Final prediction error

AIC: Akaike information criterion

SC: Schwarz information criterion

HQ: Hannan-Quinn information criterion

VAR Residual Portmanteau Tests for Autocorrelations

Null Hypothesis: No residual autocorrelations up to lag h

Date: 11/09/22 Time: 18:42

Sample: 1974 2018

Included observations: 44

Lags	Q-Stat	Prob.*	Adj Q-Stat	Prob.*	df
1	2.559648	---	2.619175	---	---
2	6.534559	0.1626	6.783367	0.1478	4

*Test is valid only for lags larger than the VAR lag order.

df is degrees of freedom for (approximate) chi-square distribution

VAR Residual Serial Correlation LM Tests

Date: 11/09/22 Time: 18:42

Sample: 1974 2018

Included observations: 44

Null hypothesis: No serial correlation at lag h

Lag	LRE* stat	df	Prob.	Rao F-stat	df	Prob.
1	3.784071	4	0.4360	0.957056	(4, 76.0)	0.4361
2	4.153814	4	0.3856	1.053117	(4, 76.0)	0.3857

Null hypothesis: No serial correlation at lags 1 to h

Lag	LRE* stat	df	Prob.	Rao F-stat	df	Prob.
1	3.784071	4	0.4360	0.957056	(4, 76.0)	0.4361
2	8.307898	8	0.4040	1.054261	(8, 72.0)	0.4046

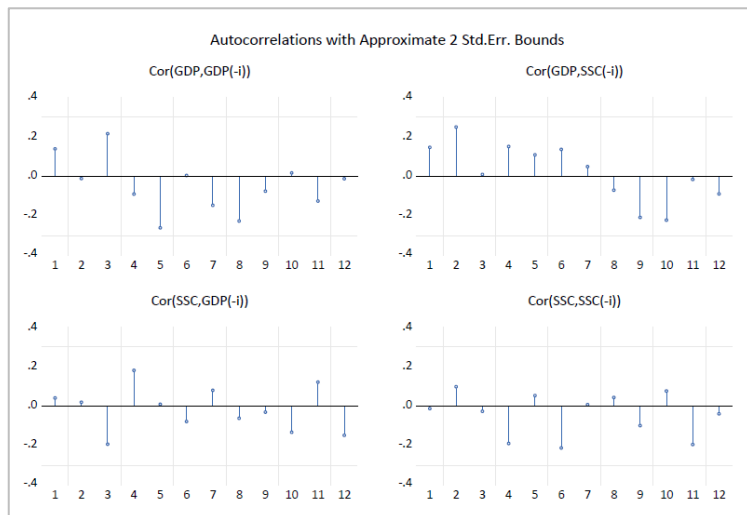
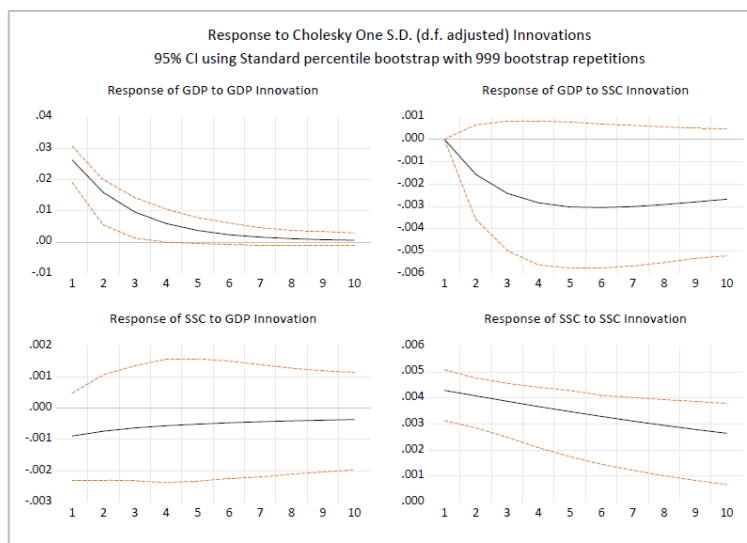
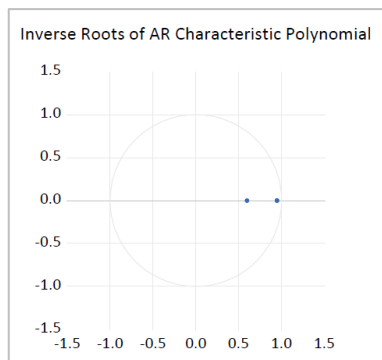
*Edgeworth expansion corrected likelihood ratio statistic

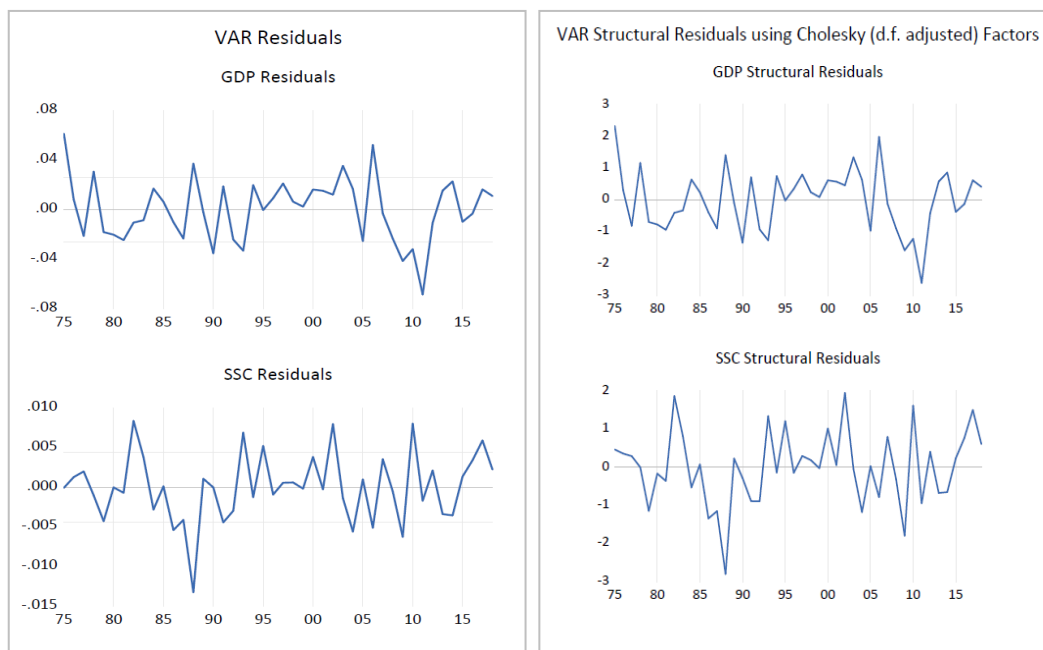
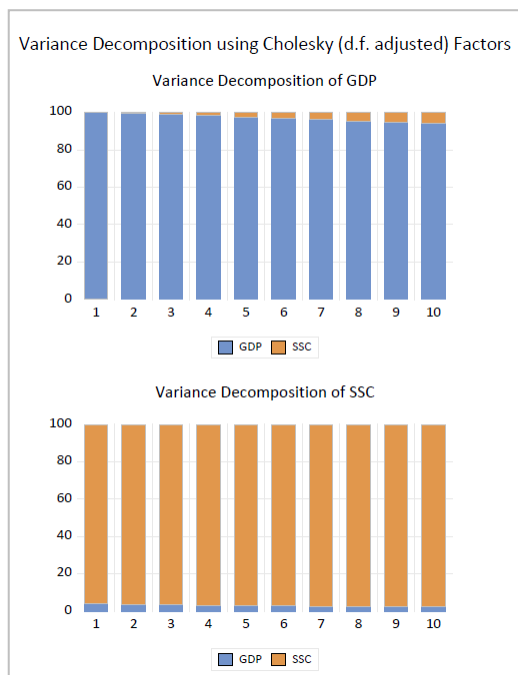
VAR Residual Normality Tests Orthogonalization: Cholesky (Lutkepohl) Null Hypothesis: Residuals are multivariate normal Date: 11/09/22 Time: 18:42 Sample: 1974 2018 Included observations: 44				
Component	Skewness	Chi-sq	df	Prob.*
1	-0.065945	0.031891	1	0.8583
2	-0.232345	0.395884	1	0.5292
Joint		0.427774	2	0.8074
Component	Kurtosis	Chi-sq	df	Prob.
1	3.203238	0.075727	1	0.7832
2	3.386428	0.273766	1	0.6008
Joint		0.349493	2	0.8397
Component	Jarque-Bera	df	Prob.	
1	0.107618	2	0.9476	
2	0.669649	2	0.7155	
Joint	0.777267	4	0.9415	
*Approximate p-values do not account for coefficient estimation				

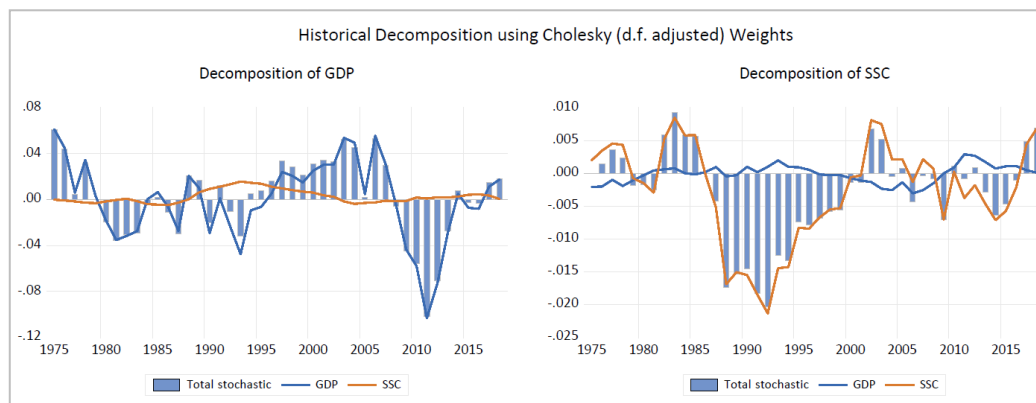
VAR Residual Heteroskedasticity Tests (Levels and Squares)					
Date: 11/09/22 Time: 18:42					
Sample: 1974 2018					
Included observations: 44					
Joint test:					
Chi-sq	df	Prob.			
18.53807	12	0.1003			
Individual components:					
Dependent	R-squared	F(4,39)	Prob.	Chi-sq(4)	Prob.
res1*res1	0.297124	4.121575	0.0070	13.07345	0.0109
res2*res2	0.101315	1.099180	0.3706	4.457840	0.3476
res2*res1	0.069964	0.733470	0.5747	3.078437	0.5448

VAR Residual Heteroskedasticity Tests (Includes Cross Terms)					
Date: 11/09/22 Time: 18:43					
Sample: 1974 2018					
Included observations: 44					
Joint test:					
Chi-sq	df	Prob.			
20.15905	15	0.1659			
Individual components:					
Dependent	R-squared	F(5,38)	Prob.	Chi-sq(5)	Prob.
res1*res1	0.320437	3.583660	0.0094	14.09923	0.0150
res2*res2	0.104094	0.883031	0.5019	4.580128	0.4692
res2*res1	0.085314	0.708865	0.6204	3.753829	0.5854

Date: 11/09/22 Time: 18:43 Sample (adjusted): 1977 2018 Included observations: 42 after adjustments Trend assumption: Linear deterministic trend Series: GDP SSC Lags interval (in first differences): 1 to 2				
Unrestricted Cointegration Rank Test (Trace)				
Hypothesized No. of CE(s)	Eigenvalue	Trace Statistic	0.05 Critical Value	Prob.**
None	0.143194	7.706376	15.49471	0.4973
At most 1	0.028527	1.215552	3.841465	0.2702
Trace test indicates no cointegration at the 0.05 level * denotes rejection of the hypothesis at the 0.05 level **MacKinnon-Haug-Michelis (1999) p-values				
Unrestricted Cointegration Rank Test (Maximum Eigenvalue)				
Hypothesized No. of CE(s)	Eigenvalue	Max-Eigen Statistic	0.05 Critical Value	Prob.**
None	0.143194	6.490824	14.26460	0.5511
At most 1	0.028527	1.215552	3.841465	0.2702
Max-eigenvalue test indicates no cointegration at the 0.05 level * denotes rejection of the hypothesis at the 0.05 level **MacKinnon-Haug-Michelis (1999) p-values				
Unrestricted Cointegrating Coefficients (normalized by b*S11*b=I):				
GDP	SSC			
-35.22662	-13.54866			
-6.819888	-66.39533			
Unrestricted Adjustment Coefficients (alpha):				
D(GDP)	0.008832	-0.001031		
D(SSC)	8.31E-06	0.000733		
1 Cointegrating Equation(s): Log likelihood 270.3132				
Normalized cointegrating coefficients (standard error in parentheses)				
GDP	SSC			
1.000000	0.384614			
	(0.72460)			
Adjustment coefficients (standard error in parentheses)				
D(GDP)	-0.311118			
	(0.13181)			
D(SSC)	-0.000293			
	(0.02547)			







Note

(1) See Asimakopoulos (2025)

References

- Asimakopoulos, P. (2025a). Driving tax administration excellence: An innovative balanced scorecard perspective. *SPOUDAI Journal of Economics and Business*, 75(1-2), 34–47.
- Asimakopoulos, P. (2025b). Evolution of Greek tax system. *Journal of Economics and Econometrics*, 68(3), 101–116. (also, EERI RP 2025/04, Economics and Econometrics Research Institute, Brussels)
- Asimakopoulos, P. (2025c). Greek tax policy and macroeconomic framework after the political transition: A systematic approach. *Theoretical and Applied Economics*, 0(2(643), S), 261–280.
- Asimakopoulos, P. (2025d). Macroeconomic impact of tax changes: The case of Greece from 1974 to 2018. *Journal of Economics and Econometrics*, 68(2), 45–68. (also EERI RP 2025/02, Economics and Econometrics Research Institute, Brussels)
- Asimakopoulos, P. (2025e). Measure what matters: A balance scorecard for tax administrations. *Global Journal of Accounting and Economy Research*, 6(1), 63–72.
- Asimakopoulos, P. (2025f). Tax and macroeconomic framework: The case of Greece. *Theoretical and Applied Economics*, 32(1(642)), 237–250.
- Asimakopoulos, P. (2025g). The impact of property tax changes in support of sustainable tax revenue in Greece. *Acta Scientiarum Polonorum. Oeconomia*, 24(1), 5–14.

Predicting customer churn using Machine Learning: models and core approaches

Mohammad Sadegh SALEM

Bucharest University of Economic Studies, Romania
sadeghsalem@gmail.com

Abstract. *Customer retention is fundamental to the growth of customer-centric businesses, especially in highly competitive subscription-based markets where customers can easily switch providers. High customer churn—the tendency of customers to stop engaging with a brand—can lead to significant material losses, reputational damage, and increased customer acquisition costs. This article reviews key strategies for proactive churn management using Machine Learning (ML) models. The findings indicate that integrating behavioral features with segmented ML models significantly enhances predictive accuracy compared to baseline models. The study discusses the essential steps in the ML workflow for churn prediction, including data collection, feature engineering, and customer segmentation. By identifying the research gap in real-time intervention strategies, this paper provides a framework for aligning ML outputs with marketing analytics.*

Keywords: machine learning, economic sciences, cloud computing, customer churn, reinforcement learning.

JEL Classification: C63, G17.

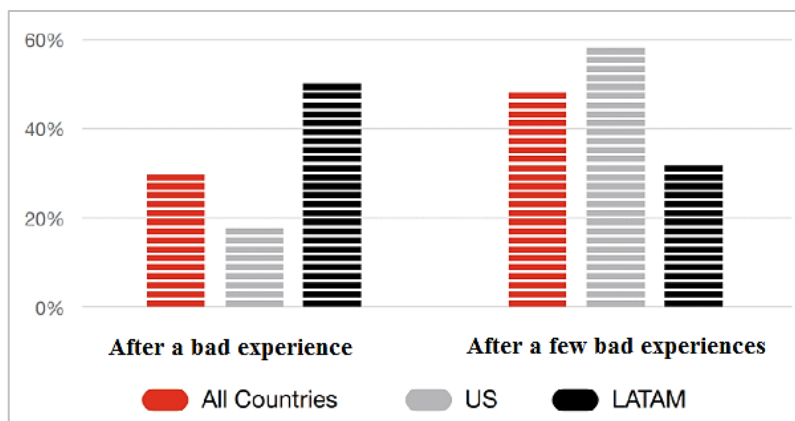
1. Introduction

Customer churn represents a major challenge for businesses, particularly those operating on a subscription or recurring service model. In these models, customers have the opportunity to reconsider their commitment at the end of each billing cycle, making them susceptible to more attractive or beneficial offers. The churn rate, a key health indicator for subscription-based businesses, is calculated as the percentage of customers who discontinue service over a specific period. For example, acquiring 150 new customers and losing three in a month results in a 2% monthly churn rate.

Negative customer experience is a primary driver of churn. Data from PwC (2018) shows that multiple bad experiences can lead 59% of US respondents to abandon a brand, while 17% will do so after only a single bad experience. Figure 1 illustrates these trends, including for LATAM (Latin America). Furthermore, high churn rates are linked to lost revenue and increased customer acquisition costs, which can be up to five times higher than retention costs (Verbeke et al., 2020). Churn also impacts future sales efforts, as dissatisfied buyers are vocal on social and review sites, leading to negative word-of-mouth that erodes trust in the company.

The Impact of Customer Churn on Business Performance: Customer churn is a detrimental occurrence for any business. The precise manner in which it influences long-term company performance warrants careful examination.

Figure 1: Customer churn chart after one or more bad experiences



Source: PwC, 2018

The General Manager of Hub Service at HubSpot emphasizes the critical nature of not underestimating the effect of even a small percentage of churn. Given that acquiring a new customer is significantly more costly than retaining an existing one, businesses with high churn rates quickly face a financial deficit. This occurs because they must allocate progressively more resources toward new customer acquisition efforts. (HubSpot, 2022)

Numerous surveys focusing on the costs of customer acquisition versus retention are readily available online. Based on one set of statistics, obtaining a new customer may cost up to five times more than retaining a current one. High churn rates are intrinsically linked

to lost revenue and elevated customer acquisition costs (CAC). Furthermore, churn rates play a critical role in an organization's potential for growth.

The Role of Word-of-Mouth and Trust in Churn: Contemporary consumers are increasingly willing to share their experiences with vendors through various channels, including review sites and social media platforms. Research by HubSpot indicates that 49% of buyers have shared an experience they had with a company on social media. (Kumar & Reinartz, 2024).

In an environment where trust in businesses is diminishing, word-of-mouth promotion plays a more significant role than ever in the purchasing process. This shift is substantiated by data revealing that:

- 55% of buyers no longer trust the companies they purchase from as much as they once did.
- 65% do not trust company press releases.
- 69% do not trust advertising.
- 71% do not trust sponsored social media advertisements.

This highlights how customer churn can be exacerbated by negative public sentiment and how lost trust further compounds the financial and reputational impact of a high churn rate.

Therefore, it can be concluded that companies with high churn rates not only fail in their relationships with former customers, but also harm their future purchasing efforts by creating negative word of mouth about their products.

High churn rates are intrinsically linked to lost revenue and elevated customer acquisition costs (CAC). Contemporary consumers are increasingly willing to share their experiences through social media, where 49% of buyers have shared a company experience (HubSpot, 2022). In an environment where 55% of buyers no longer trust the companies they purchase from, word-of-mouth plays a critical role in growth (Kumar & Reinartz, 2024).

1.1. Research Questions and Hypotheses

This study addresses the following research questions:

RQ1: How can ML-driven behavioral patterns predict churn in subscription-based models?

RQ2: What is the impact of customer segmentation on the precision of retention strategies?

Hypothesis (H1): Proactive ML-based monitoring reduces CAC by enabling timely interventions.

The central hypothesis (H1) posits that proactive ML-based monitoring reduces customer acquisition costs (CAC) by enabling timely interventions.

2. Literature Review and methodology

This study employs a library research methodology to explore the application of machine learning algorithms in economic sciences, particularly in the context of cloud computing. The methodology involves a systematic review and analysis of existing literature, which includes academic papers, books, reports, and relevant online resources. The primary focus

of the research is to identify the computational challenges in economic sciences and evaluate how cloud-based machine learning algorithms can address these challenges more effectively than traditional methods. The key steps in the research design are:

- **Literature Review:** A comprehensive review of literature on machine learning, cloud computing, and their applications in economic sciences was conducted. This includes peer-reviewed journals, conference papers, books, and reputable online sources. The literature review serves as the foundation for understanding the state-of-the-art in these fields and identifying gaps that the current study aims to address.
- **Thematic Analysis:** The literature is analyzed thematically to identify the main computational problems in economic sciences, the types of machine learning algorithms used, and the role of cloud computing in enhancing these algorithms. Themes such as scalability, computational efficiency, and cost-effectiveness are explored.

Proactive Churn Management and Predictive Analytics

To address churn proactively, businesses are increasingly adopting predictive analytics. This form of analysis involves forecasting the probability of future outcomes, such as a customer churning, by analyzing current and historical data (Hadden et al., 2021). Predictive analytics uses statistical techniques, including data mining and Machine Learning (ML).

Figure 2: Machine learning model



Source: Braze n.d., 2025

ML systems are designed to find patterns in data and learn from them without explicit programming. In the context of churn, these systems identify behavioral patterns—such as a decrease in customer activity—that indicate declining satisfaction. By detecting these patterns, ML algorithms can signal potential churners among the current customer base, allowing the organization to intervene actively and understand the root causes of the customer's issues before it is too late. Identifying high-risk customers enables organizations to engage proactively and understand pain points (Ascarza et al., 2020).

Companies that consistently monitor how individuals interact with their products, encourage feedback, and resolve issues promptly are significantly better positioned to maintain mutually beneficial customer relationships and enhance retention.

Consider an organization that systematically collects customer data over time. This data can be leveraged to identify patterns of potential behavioral deviation, segment high-risk customers, and implement targeted interventions to rebuild trust. Organizations adopting a proactive approach to churn management typically utilize predictive analytics.

The Role of Predictive Analytics and Machine Learning: Predictive analytics is a type of analysis that involves forecasting the probability of future outcomes, events, or values by examining current and historical data. This methodology employs various statistical techniques, notably Data Mining (for pattern recognition) and Machine Learning (ML).

The core feature of Machine Learning is its ability to construct systems capable of finding patterns within data and learning from them without explicit programming. In the context of predicting customer churn, these patterns often manifest as online behavioral features that signal a decline in customer satisfaction with the company's services or products (Von Luxburg, U, 2007).

Identifying customers at risk of churning facilitates timely, pre-emptive action. Machine Learning methods are critically important for proactive churn management, as algorithms excel at identifying potential churners. They reveal common behavioral patterns among customers who have already departed the company. Subsequently, ML algorithms continuously audit the behavior of current customers against these established patterns, flagging potential churn risks when detected.

Strategic Application of Predictive Modeling: Businesses can employ ML for predictive analysis to understand which current users are not fully satisfied with their service, allowing them to address issues before it is too late. Identifying high-risk customers enables an organization to:

- Engage proactively with these customers.
- Understand their pain points and reasons for dissatisfaction.
- Formulate a long-term plan focused on helping the customer re-perceive the value of the purchased services.

The use cases for predictive churn modeling extend beyond active customer engagement and selecting effective retention measures. ML-driven software allows business managers to determine which specific customers to contact. In essence, employees can be confident they are speaking with the right customers at the right time (Iordanis, K., and Landman, J., 2021).

Sales, customer success, and marketing teams can also leverage the insights from this data analysis to coordinate their efforts. For example:

- If a customer shows signs of churn risk, it is likely an inopportune time for a sales-driven upsell attempt regarding additional services.
- Instead, this interaction should be handled by the Customer Success Management (CSM) department to help the customer re-engage and recognize the value of the products they already possess.
- Similarly, the Marketing department can tailor its interactions with customers based on current churn risk indicators.

- Recent empirical studies demonstrate that churn prediction is most effective when linked to Customer Lifetime Value (CLV) models. For instance, companies using random forest algorithms have reported a 15-20% increase in retention by focusing on high-value segments."

Utilizing Machine Learning for customer data analysis provides the crucial insights necessary to strengthen this coordinated, predictive retention strategy.

3. The Machine Learning Workflow for Churn Prediction

To initiate the process of customer churn rate prediction, one must understand the necessary data and the typical implementation steps. As with any Machine Learning (ML) endeavor, data science professionals must first focus on data acquisition (Kerenidis, I., Luongo, A., and Prakash, A , 2020).

Data Requirements and Preparation: The research objectives dictate the specific data requirements. Researchers must define which data sets are relevant to the prediction goal. These data sets are subsequently prepared, pre-processed, and transformed into a format suitable for building ML models.

Model Development and Deployment: A critical phase of the work involves selecting appropriate training methods, fine-tuning the models, and ultimately choosing the best-performing model. The model selected must exhibit the highest predictive accuracy. Once validated, the chosen model can then be successfully deployed into a production environment.

The general workflow data scientists follow to build ML-powered systems capable of predicting customer churn typically involves the following key stages:

- **Data Collection and Definition:** Identifying and gathering relevant historical customer data (e.g., usage logs, support tickets, billing information, demographic data).
- **Data Pre-processing and Feature Engineering:** Cleaning the data, handling missing values, scaling features, and creating new, predictive features from the raw data (De Caigny et al., 2020).
- **Model Selection and Training:** Choosing appropriate ML algorithms (e.g., classification algorithms like Logistic Regression, Decision Trees, or Random Forests) and training them on the historical data.
- **Model Evaluation and Fine-Tuning:** Assessing the model's performance using metrics (e.g., accuracy, precision, recall, F1-score) and optimizing hyperparameters.
- **Model Deployment (Go-to-Production):** Integrating the final, validated model into the business's operational systems to continuously score current customers for churn risk.
- **Monitoring and Maintenance:** Continuously monitoring the model's performance in a live environment and updating it as customer behavior or business conditions change (model drift).

This structured approach ensures the development of a robust and actionable predictive model for proactive churn management.

3.1. Problem Formulation: Classification vs. Regression

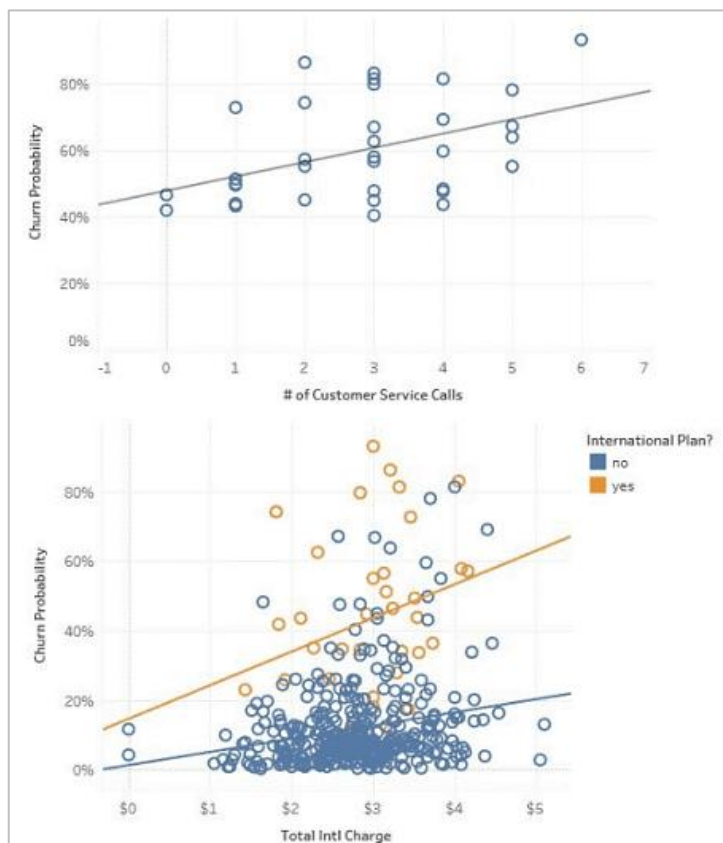
The initial step involves defining the problem and the desired insights, which dictates the type of ML task: classification or regression.

Classification: The goal is to determine which category or class the customer data belongs to. In churn prediction, this addresses binary questions such as: "Will this customer churn or not?" or "Will the customer renew their subscription?"

Regression: This involves estimating the relationship between a target variable (e.g., churn probability) and other data values, yielding a continuous, numerical result. Regression can predict the specific time period in which a customer is likely to churn or provide a numerical churn probability per customer (Fader & Hardie, 2022).

The figure below is an example of a logistic regression used to predict the probability of churn in telecommunications by Science Data Towards. Here, the figure 3 shows how the number of service calls and the use of international plans are related to churn.

Figure 3: The relationship between the number of service calls and the use of international plans with churn



Source: RapidMiner. n.d, 2025

3.2. Data Collection and Preparation

Data Sources and Features: The quality of the received insights directly depends on the quality of the data. Essential data sources for churn prediction include Customer Relationship Management (CRM) systems, analytics services (e.g., Google Analytics), and social media feedback. Data scientists then create and select a set of features (input variables) that represent different behavioral patterns related to the customer's interaction with the product or service. These features are broadly categorized into four groups:

- **Demographic Features:** Basic information (e.g., age, location, income).
- **Behavioral Features:** How the person uses the service (e.g., login frequency, session length, features used, monetary value).
- **Support Features:** Customer interactions with support (e.g., number of submitted requests, satisfaction scores).
- **Contextual Features:** Other contextual information.

Feature Selection and Segmentation: The goal of feature selection is to reduce data dimensionality by choosing only the variables most relevant to customer churn, eliminating irrelevant information. For growing companies, customer segmentation based on demographics, usage, or monetary value is also beneficial. Training separate ML models on segmented datasets can enhance predictive accuracy, as each segment shares common behavioral patterns. (Kolm, P.N., and Ritter, G, 2020).

4. Model Building and Evaluation

In the modeling phase, data scientists typically train, tune, and test multiple models to identify the one that detects potential churners with the highest accuracy on the training data.

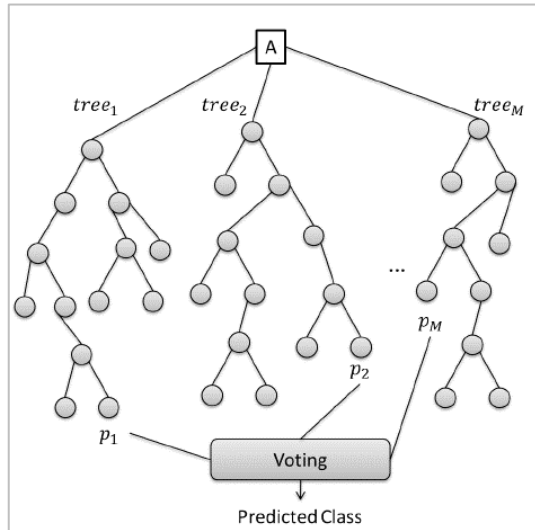
4.1. Classic Machine Learning Models

While complex models like XGBoost, LightGBM, and CatBoost can be evaluated, classic ML algorithms are often used as a baseline. These include:

Logistic Regression: Used for binary classification problems, it predicts the probability of an event (e.g., churn) occurring by measuring the relationship between dependent and independent variables.

Decision Trees: A supervised learning algorithm that splits data into homogeneous subsets based on the most significant differentiators in the input variables. Decision trees are easily interpretable and require less data preparation but can be unstable.

Random Forest: An ensemble learning method that utilizes multiple decision trees to achieve higher predictive accuracy and model stability. It operates by having each tree classify a data sample, and the forest selects the classification with the most votes (Verbeke et al., 2020). The figure 3 shows Random Forest algorithm in Machine learning model.

Figure 4: Machine learning model (Random Forest)

Source: ResearchGate n.d, 2025

4.2. Deployment and Monitoring

The final selected model is then integrated into existing software or a new application. However, churn prediction is an iterative, continuous process. Data scientists must continually monitor the model's accuracy levels and adapt or revise the features as needed to maintain the desired level of predictive performance.

5. Conclusion

Predicting customer churn with Machine Learning is a vital component of a comprehensive retention strategy. The ability to identify dissatisfied customers allows businesses to understand product weaknesses, pricing issues, and customer expectations, enabling proactive resolution. By systematically applying the ML workflow—from defining the problem as a classification or regression task, to selecting the most relevant behavioral features, segmenting the customer base, and continuously monitoring the deployed model—organizations can significantly reduce churn and foster sustainable growth.

In conclusion, this study confirms H1, showing that ML workflows provide a robust framework for reducing churn. Addressing RQ1 and RQ2, the integration of behavioral segmentation allows marketing teams to transition from reactive to proactive engagement. The primary research gap identified is the need for more dynamic models that account for shifting consumer sentiment in real-time."

References

- Braze (n.d.) (2025). Machine learning model for churn prediction [Infographic]. Braze Documentation. Retrieved from <<https://www.braze.com/>>
- RapidMiner (n.d.) (2025). Customer churn probability analysis via model simulator [Graph]. Altair RapidMiner Documentation. Retrieved from <<https://rapidminer.com/>>
- Von Luxburg, U. (2007). A Tutorial on Spectral Clustering. *Statistics and computing*, 17(4), 395-416.
- Kumar, V., and Reinartz, W. (2024). *Customer Relationship Management: Concept, Strategy, and Tools*. Springer Nature.
- PwC (2018). *Experience is everything: Get it right*. Global Consumer Insights Survey. <<https://www.pwc.com/us/en/services/consulting/library/consumer-intelligence-series/future-of-customer-experience.html>>
- Verbeke, W., Baesens, B., and Bravo, C. (2020). *Profit-Driven Business Analytics: A Practitioner's Guide to Predicting and Optimizing Outcomes*. Wiley.
- Von Luxburg, U. (2007). A Tutorial on Spectral Clustering. *Statistics and Computing*, 17(4).
- Culkin, R., and Das, S.R. (2017). Machine Learning in Finance. *Journal of Investment Management*.
- ResearchGate (n.d.) (2025). General architecture of a Random Forest classifier with voting mechanism [Diagram]. Retrieved from <<https://www.researchgate.net/>>
- Iordanis, K., and Landman, J. (2021). Quantum Spectral Clustering. *Physical Review A*, 103(4).
- Kerenidis, I., Luongo, A., and Prakash, A. (2020). Quantum Expectation-Maximization for Gaussian Mixture Models. *Proceedings of International Conference on Machine Learning (ICML)*.
- Kolm, P.N., and Ritter, G. (2020). Modern Perspectives on Reinforcement Learning in Finance. *The Journal of Machine Learning in Finance*, 1(1).
- Ascarza, E., Netzer, O., and Hardie, B.G. (2020). Some Customers Would Rather Leave Without Saying Goodbye. *Marketing Science*.
- De Caigny, A., Coussement, K., and De Bock, K.W. (2020). A new hybrid classification algorithm for customer churn prediction based on logistic regression and decision trees. *European Journal of Operational Research*.
- Fader, P.S., and Hardie, B.G. (2022). *The Customer Centricity Playbook*. Wharton School Press.
- Hadden, J., Tiwari, A., Roy, R., and Ruta, D. (2021). Computer assisted customer churn management: State-of-the-art and future trends. *Journal of Data Science and Analytics*.
- HubSpot Research (2022). *The State of Customer Service Report*.

Have we truly transitioned from industrial to service-based economies? An analysis of the economic sectoral distribution

Nora Luz QUIROZ SANTIAGO

Universidad Autónoma de Aguascalientes (UAA), México
nora.quiroz@edu.uaa.mx

Abstract. *Since the mid-20th century, it has been argued that countries would transition from industrial economies to service-based economies. This study assesses this transition by analyzing the sectoral composition of GDP in 2024 across 205 countries and territories, with a multi-scale analysis of the world, the Americas, and Mexico. Results show that services are the main sector in over 90% of countries, although development levels remain highly unequal. In Mexico, services have increased while industry has relatively declined, suggesting relative deindustrialization. Overall, although the transition to service economies is confirmed, it masks unequal structural trajectories and dynamics between countries.*

Keywords: sectoral composition, tertiarization, GDP, economic structure, México.

JEL Classification: F63, O10, L16.

1. Introduction

The three-sector theory, also known as tripartite economics, posits that economies are structured around three main types of activities: a) those related to agriculture and the direct extraction of natural resources (primary sector), b) those focused on transforming raw materials into tangible intermediate or final goods, such as manufacturing industries (secondary sector), and c) those linked to the provision of services, whose intangible nature helps to facilitate, organize, or complement both production and consumption (tertiary sector). This classification was emphasized by Clark (1940) in his treatise on the conditions of economic progress, laying the groundwork for understanding the sectoral composition of modern economies.

Beyond this tripartite division, several authors have pointed out the importance of analyzing not only the existence of these sectors, but also their relative weight and their evolution over time. In this sense, the three-sector growth hypothesis holds that economic development entails structural change and a productive transition, in which the participation of the primary sector tends to decrease, while the secondary and, subsequently, the tertiary sectors gain greater relevance (Krüger, 2008; Sonis, Azzoni and Hewings, 2008). It was Bell (2006) who further developed this idea, pointing out that in the pursuit of advanced economies, there would be this productive transition toward services, a sector that would become predominant over the others.

Both the distribution and the sectoral predominance are not only structural characteristics of nations' economic performance (Krüger, 2008), but they also play a role in shaping multiple dimensions of economic and social development, such as territorial specialization, interaction between regions and countries, sociocultural dynamics and urban mobility (Quiroz Santiago, Sánchez-Zárate and Santiago, 2025). They can even affect individual lives, particularly in the work and family spheres.

However, given that more than eight decades have passed since the tripartite economic model was proposed, a fundamental question arises: have we truly transitioned from industrial economies to advanced service-based economies? In this regard, the objective of this research is to assess whether the transition from industrial to service economies has been consolidated and to what extent this process is homogeneous or differentiated across countries and regions.

Based on the 2024 Gross Domestic Product (GDP) by economic sector, the study examines the distribution across the primary, secondary, and tertiary sectors (CIA, 2024b; 2024a). A multi-scale analytical approach is employed, encompassing three levels of analysis: the global level, covering 205 countries and territories; the Americas; and Mexico, where a three-decade longitudinal analysis examines the sectoral transition at the national level, complemented by sector analysis of its federal entities for 2024.

The study's main contribution to the literature is challenging the widely accepted narrative of a linear transition from industrial economies to advanced service-based economies and post-industrial societies. The findings indicate that, while the service sector is statistically dominant, this does not imply a homogeneous structural replacement or the disappearance of industrial bases. Rather, it reveals hybrid, differentiated, and territorially unequal trajectories.

The document is structured into sections: first, the theoretical framework; then, the methodology; next, the results; and finally, the conclusions.

2. Theoretical framework

2.1. Tripartite economy

The three-sector theory, also known as tripartite economics, was emphasized by Colin Clark (1940) in his treatise on the conditions of economic progress, laying the groundwork for understanding the sectoral composition of economies. It posits that economies are organized around three types of activities: agriculture, industry, and services. This classification has evolved, with these activities now referred to as the primary, secondary, and tertiary sectors. For example, in 2024, Mexico's economy was composed of 3.4% in the primary sector, 33.4% in the secondary sector, and 63.2% in the tertiary sector (INEGI, 2024a). To provide a clearer understanding of this classification, the economic activities comprising each sector are detailed below:

- Primary activities (agriculture) are associated with agriculture and the direct extraction of natural resources, without these having undergone complex industrial transformations. This also includes animal husbandry and farming, forestry, fishing, and hunting. Its central function is to provide basic inputs and raw materials, which constitute the material basis for the development of other economic activities (INEGI, 2024a).
- Secondary activities (industry) are dedicated to the physical, chemical, or mechanical transformation of raw materials into tangible intermediate or final goods. They are generally associated with manufacturing industries, but in Mexico's classification also include construction, mining, certain forms of energy production, and the supply of both water and natural gas (INEGI, 2024a). Through these activities, economic value is added to natural resources, allowing their adaptation for consumption, productive use, or infrastructure.
- Tertiary activities (services) are linked to the provision of services, whose intangible nature contributes to facilitating, organizing, or complementing both production and consumption; acting as a link between the different economic sectors and society. In the SCIAN, the adaptation for Mexico of the North American Industry Classification System (NAICS), the tertiary sector includes trade (wholesale and retail), a wide range of services (mass media information, financial, real estate and accommodation, professional, business support, corporate, waste management, remediation, educational, health and social assistance, recreational and sports, food and beverage preparation), legislative, governmental and other activities (INEGI, 2023).

These three sectors constitute the main components of an economy's sectoral structure and contribute to its gross domestic product (GDP), which measures "the monetary value of the final goods and services produced by an economy in a given period" (INEGI, 2017). For example, Mexico's GDP increased from 20.1 to 25.4 trillion Mexican pesos between 2010 and 2024. Until 2018, it shows sustained growth (20.1 to 24.2), stagnation in 2019 (24.1), and a decline in 2020 (22.1), likely associated with the economic effects of the COVID-19 pandemic. Subsequently, it experienced a continuous recovery that surpassed previous levels (23.4 to 25.4) (INEGI, 2024a).

Historically, GDP dates to the 17th century, when William Petty developed a basic version for fiscal purposes. The concept was formalized in the 1930s by Simon Kuznets, who presented the Kuznets curve twenty years later, which helps us understand how economic development affects income and inequality. It posits that during the transition from agrarian to industrialized economies, inequality first increases, peaks, and then declines. This inequality, including in the transition to a service-based economy, tends to increase due to the skills and opportunities gap. However, with economic maturity, education, and technological progress, this gap narrows.

GDP estimates consider four components: consumption, investment, government spending, and exports. The result can be expressed in nominal terms, when calculated at current prices, or in real terms, when adjusted for inflation. However, to understand a country's economic structure, net product taxes are also calculated alongside measures of the three economic activities. Taxes represent the sum of taxes levied on goods and services minus subsidies, thus measuring the impact of fiscal intervention on the economy.

Thus, GDP, in addition to being an indicator of the economic health and structure of countries, allows us to measure production growth or decline and business competitiveness. Furthermore, it helps guide the formulation of public policies, support investment decisions, analyze—with certain limitations—its implications for social welfare, and make international comparisons (CIA, 2024a; 2024b).

2.2. Problems with sectoral classification and GDP

Although the tripartite classification of the economy has been widely used in economic and territorial studies, due to its analytical clarity, international comparability—albeit with its due equivalences—and ability to offer a synthetic view of the productive structure; nevertheless, this same structure has been subject to theoretical and empirical questioning, mainly due to its limited capacity to capture the functional and territorial complexity of contemporary economies.

In the economic sphere, criticisms focus on the limited disaggregation of diverse activities and the insufficient incorporation of changes inherent to contemporary economies. As for GDP, the criticisms are varied: it fails to capture real well-being, tends to prioritize economic growth over environmental issues, and overlooks inequality by relying on averages that assume a uniform distribution and can mask significant gaps. Furthermore, it does not account for informal work, domestic labor, or care work.

Regarding the service sector, Schafran et al. (2018) have noted that grouping services into a single category can obscure substantial differences among service types, particularly those related to knowledge and innovation. They also note that this typology does not explicitly include activities associated with urbanization and control processes.

Furthermore, the technological and cultural changes of recent decades, associated with globalization, the reconfiguration of production chains, and digitalization, have significantly transformed the nature of economic activities. As a result of these transformations, there is a growing shift from economies previously based on industry to knowledge-based models (Cervelló-Royo et al., 2026).

This has led to proposals for a more detailed breakdown of the tertiary sector and even the incorporation of new analytical categories, such as quaternary activities (Selstad, 1990), associated with knowledge, information, and advanced services (Schafran et al., 2018). These proposals seek to address the growing functional heterogeneity of economic activities and the classification of the resulting jobs. They also incorporate other indices, such as the Human Development Index (HDI) and the Real Progress Index (RPI).

2.3. Economic transition, service sector and advanced economies

The three-sector growth hypothesis argues that as economies develop, they undergo structural changes that lead to productive transitions. The foundations of this process can be traced back to the 17th century with Petty's Law, which stated that incomes were higher in manufacturing than in agriculture, and even higher in goods than in manufacturing. (Petty, 1676 in Hull, 1899). Although there was no explicit formulation of the three-sector economy at that time, structural transformation and cost reduction were already recognized.

The law was revisited and refined, resulting in the Petty-Clark Act, which posited that this transition was also reflected in the structure of employment, with a shift of the workforce from agriculture to industry and, subsequently, to services (Clark, 1940). Along these lines, Fourastié (1949) emphasized that this productive transition, coupled with technological progress and increased labor productivity, would bring various benefits, such as improvements in quality of life, educational and cultural development, increased job skills, and reduced unemployment (Sonis, Azzoni and Hewings, 2008).

Bell (2006) further developed this idea by suggesting that, on their quest to become advanced economies, countries might experience a transition towards the dominance of the service sector. From this perspective, it is hypothesized that the least developed countries with an agricultural or pre-industrial base would become developing countries with an industrial base and, eventually, developed countries with a predominance of the service sector.

3. Methodology

To answer the central question: have we truly transitioned from industrial economies to advanced service-based economies? A quantitative and descriptive-structural approach has been adopted, focusing on the sectoral percentage distribution of global economies. The aim is to assess whether the transition from industrial to service-based economies has taken hold and whether this process is homogeneous or differentiated across countries by their respective levels of development.

The analysis is developed from a multi-scale perspective, integrating information at the global, American, and Mexican scales from two databases. First, GDP figures for 2024 are used, along with the percentage shares of the three economic sectors, across 205 countries and territories⁽¹⁾. These figures were obtained from the World Bank and published by CIA (2024b, 2024a). In this tripartite classification, agriculture also includes fishing, hunting, and forestry; industry encompasses mining, manufacturing, power generation, and construction; while services include government activities, communications, transportation, finance, and others (CIA, 2024a).

Secondly, for the specific study of Mexico, the sectoral distribution from 1993 to 2024 is analyzed longitudinally, as well as the structure by Mexican federal entity for the year 2024, based on the GDP information obtained by the INEGI (2025), which uses a classification composed of primary, secondary, tertiary activities, and taxes.

The results of these analyses are presented in the following section.

4. Results

4.1. Sectors on a global scale

In 2024, global GDP reached \$173.163 trillion, with a predominantly tertiary-sectoral distribution⁽²⁾, in which services accounted for 66.2%, followed by industry (26%) and, to a lesser extent, agriculture (4%). This configuration demonstrates a structural transition towards service-oriented economies worldwide, since, of the 205 countries studied, services are the main contributor to GDP in 191 (93.17%). When investigating the level of development of these countries, it is discovered that this is not a homogeneous group, since approximately 1 in 4 are developed countries (23.56%, 45), approximately 1 in 2 are developing countries (48.17%, 92) —Mexico is in this category, ranking 13th in world GDP in 2024—, approximately 1 in 5 are least developed countries (19.9%, 38) and approximately 1.6 in 10 are special territories (8.38%, 16).

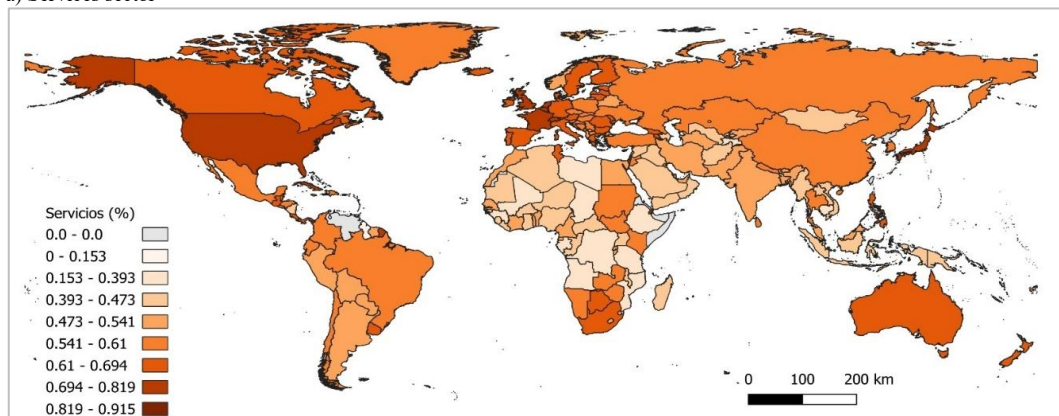
The highest percentages of economic activity attributable to the services sector are found in Bermuda (91.50%), Hong Kong (91.40%), and Macau (91.40%). When analyzed by region (see Map 1.a), in Europe, the following stand out: United Kingdom (72.80%), France (70.40%), Netherlands (70.30%), Spain (69.10%), Italy (65.60%), Germany (63.90%), among many others. In Africa: São Tomé and Príncipe (76.6%), Djibouti (75.5%), Botswana (63.50%), South Africa (62.70%), Egypt (48.90%), Nigeria (47%), and others. In Asia: Japan (69.8%), Philippines (63.2%), Thailand (59.2%), South Korea (58.4%), China (36.5%), and many others. In Oceania: New Zealand 67.4% and Australia 65.5%. In North America: United States 79.7% and Canada 66.40%; and in Latin America: Costa Rica 68.8%, Panama 68.8%, Uruguay 65.3%, El Salvador 61%, Brazil 59.30%, Colombia 58.2%, Mexico 58.20%, Chile 56.1%, Argentina 53.4% and many others.

Regarding industry as the main sector (see Map 1.b), only 13 of the 205 countries—equivalent to 6.34% of the total and ranked between 43rd and 140th in the world GDP ranking—have it. Of these, only three (1.46% of the total) are developed countries located in Asia (Qatar—59th in world GDP in 2024—, Kuwait, and Brunei); eight are developing countries located in Asia, Africa, and one in Latin America (Guyana); while the remaining two are least developed countries in Africa.

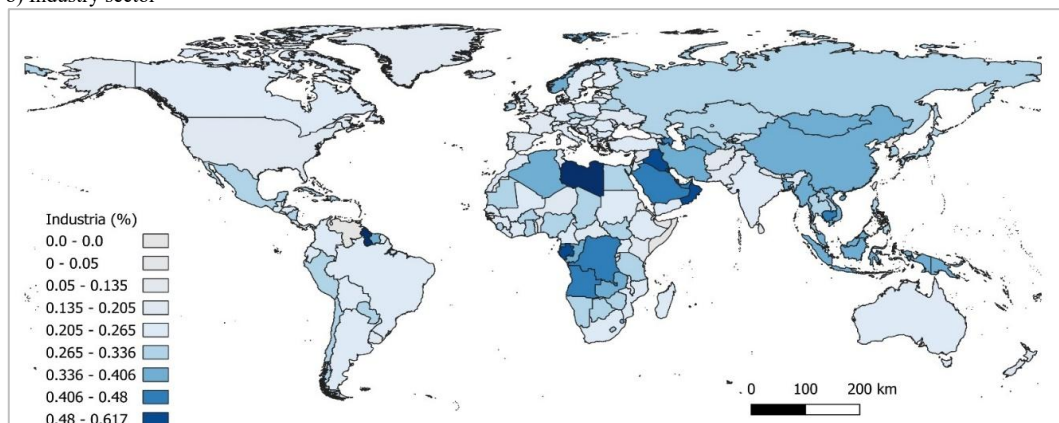
Regarding the economic dominance of agriculture (see Map 1.c), only Chad, located in Central Africa, has it. However, it is important to note that its economic distribution is quite balanced (agriculture 32.20%, industry 29.70%, and services 31.60%); this suggests the need to analyze not only the main sector but also the percentage distribution of the three sectors.

Map 1 *Percentage contribution (%) of GDP by sector*

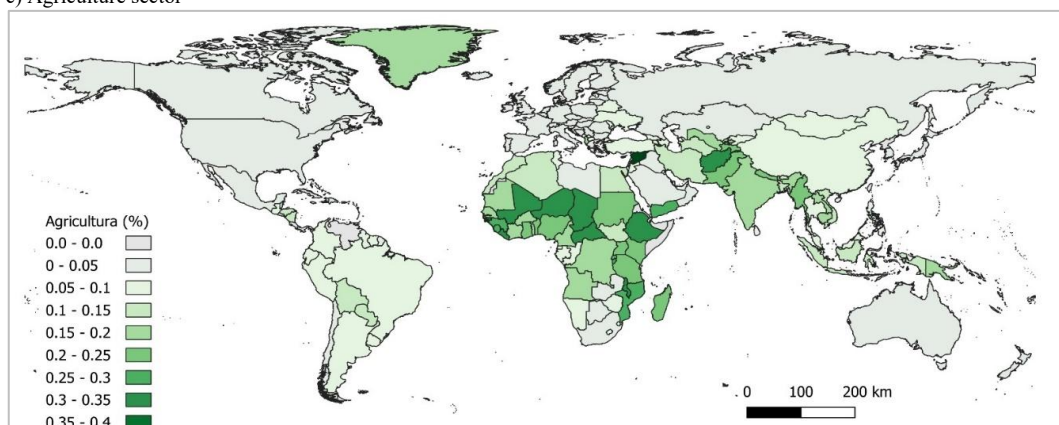
a) Services sector



b) Industry sector



c) Agriculture sector



Note: Several countries do not formally participate in the World Bank's PPP project that calculates these measures. Hence, the resulting GDP estimates are inaccurate, or figures are not reported, as is the case in Venezuela.

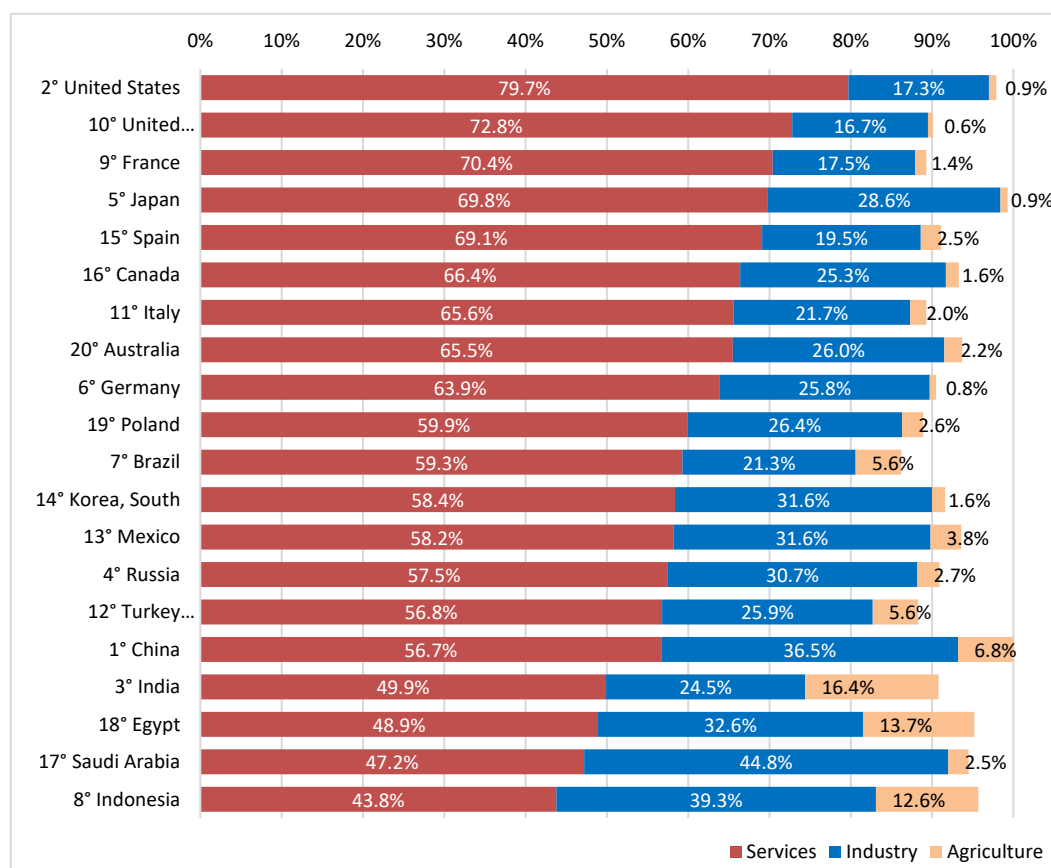
Source: Prepared by the author using information from CIA (2024b, 2024a).

4.2. Twenty countries with the highest GDP in the world in 2024

Having established the importance of identifying not only the predominant sector in GDP but also the percentage contribution of each of the three economic sectors, twenty countries with the highest GDP in 2024 will be analyzed⁽³⁾ (Figure 1). In terms of their level of development, almost half of these countries (11) are classified as developed, slightly outnumbering developing countries (9). Regarding sectoral dominance, services are the largest sector in all countries analyzed, with shares ranging from 43.8% in Indonesia (8th) to 79.7% in the United States of America (2nd).

The sectoral distribution analysis reveals that, although China ranks first globally, its share in the services sector (56.7%) is lower than that of the second-placed United States (79.7%). Similarly, it is nearly on par with Turkey (56.8%). In the industrial sector, China ranks third globally (36.5%), surpassed only by Indonesia (39.3%, second) and Saudi Arabia (44.8%, first). This highlights the importance of sectoral distribution in each country's economy.

Figure 1. Countries with the highest GDP in 2024 (descending order by services sector)



Notes: Figures may not add up to 100% because unallocated consumption has not been captured in the sector-reported data. GDP information is for 2024, except for Japan and South Korea, whose GDP data is for 2023; and for Canada, whose sector-specific information is for 2021.

Source: Prepared by the author using information from CIA (2024b, 2024a).

4.3. Sectoral distribution in the American continent

Continuing the analysis at the continental level, the focus now shifts to the Americas. As shown in Table 1, services constitute the predominant sector across all countries, with shares ranging from 46.8% to 79.7%. Guyana is the sole exception, with industry representing the largest share at 74.3%.

Table 1 Americas: Percentage distribution of GDP by economic sector

Country	Year	GDP	Agriculture (%)	Industry (%)	Services (%)	Main sector	%/100
2° United States**	2024	\$25.676 trillion	0.90%	17.30%	79.70%	Services	81.41%
7° Brazil	2024	\$4.165 trillion	5.60%	21.30%	59.30%	Services	68.79%
13° Mexico	2024	\$2.883 trillion	3.80%	31.60%	58.20%	Services	62.18%
16° Canada*	2021	\$2.341 trillion	1.60%	25.30%	66.40%	Services	71.17%
28° Argentina	2024	\$1.213 trillion	6.00%	24.00%	53.40%	Services	64.03%
31° Colombia	2024	\$978.592 billion	9.30%	23.10%	58.20%	Services	64.24%
42° Chile	2024	\$596.556 billion	3.90%	30.10%	56.10%	Services	62.26%
46° Peru	2024	\$535.911 billion	6.10%	32.20%	52.70%	Services	57.91%
64° Dominican Republic	2024	\$276.884 billion	4.50%	28.70%	59.80%	Services	64.30%
67° Ecuador	2024	\$252.728 billion	9.50%	26.50%	57.20%	Services	61.37%
70° Guatemala	2024	\$232.673 billion	9.80%	21.70%	61.80%	Services	66.24%
79° Panama	2024	\$164.484 billion	2.60%	26.30%	68.80%	Services	70.42%
85° Puerto Rico	2024	\$141.344 trillion	0.70%	48.00%	51.50%	Services	51.40%
86° Costa Rica	2024	\$138.371 trillion	3.60%	19.70%	68.80%	Services	74.70%
91° Bolivia	2023	\$120.531 trillion	13.50%	24.20%	51.10%	Services	57.55%
92° Paraguay	2024	\$112.919 trillion	10.70%	32.50%	48.70%	Services	52.99%
94° Uruguay	2024	\$108.502 trillion	6.40%	16.80%	65.30%	Services	73.79%
104° Cuba	2023	\$81.985 trillion	1.30%	27.50%	70.00%	Services	70.85%
106° El Salvador	2024	\$73.961 trillion	4.40%	22.40%	61.00%	Services	69.48%
108° Honduras	2024	\$71.297 trillion	11.20%	26.10%	58.40%	Services	61.02%
118° Guyana	2024	\$58.423 trillion	8.00%	74.30%	15.30%	Industrial	76.13%
122° Nicaragua	2024	\$52.989 trillion	14.40%	27.60%	46.80%	Services	52.70%
137° Trinidad and Tobago	2023	\$42.658 trillion	0.80%	35.00%	59.90%	Services	62.59%
145° Haiti***	2024	\$32.971 trillion	15.90%	33.40%	48.30%	Services	49.49%
148° Jamaica	2024	\$29.13 billion	9.80%	18.30%	60.30%	Services	68.21%
157° Bahamas	2024	\$14.544 trillion	0.50%	9.60%	77.20%	Services	88.43%
180° Belize	2023	\$2.657 trillion	1.90%	19.00%	69.10%	Services	76.78%

Notes: The total figures for the three sectors may not add up to 100% because unallocated consumption has not been captured in the sector-reported data. *2021 GDP, **Developed country, ***Least developed country. Several countries do not formally participate in the World Bank's PPP project that calculates these measures, so the resulting GDP estimates are not accurate. Figures for Venezuela are unavailable.

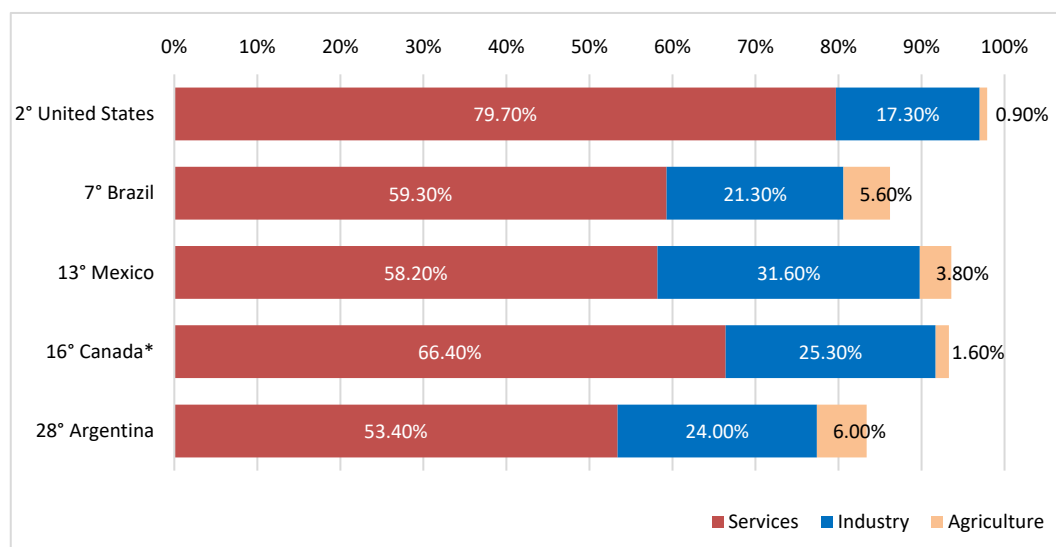
Source: Prepared by the author using information from CIA (2024b, 2024a).

Figure 2 analyzes the five largest economies in the Americas by GDP. The country at the top of the American ranking is the United States (2nd in the world), with the highest share of economic activity concentrated in the service sector (79.7%). Meanwhile, industry plays a secondary role (17.3%), while agriculture accounts for only a marginal share (0.9%). Canada, though with a slightly larger industrial presence (24%), follows the USA in terms of distribution, with a similar pattern reinforcing the fact that North America has advanced, highly service-oriented economies.

Brazil, Mexico, and Argentina occupy an intermediate position, where the service sector remains the largest (ranging from 53.4% to 59.3%) while coexisting with a significant industrial base. Mexico stands out for its high industrial share (31.6%), higher than those of Brazil (21.3%) and Argentina (24%), reflecting its role as a manufacturing economy in

the region. In other words, this profile reflects Mexico's role as an intermediate economy, with a significant industrial base (almost one-third) and a growing orientation toward tertiary activities (almost two-thirds). In contrast, Argentina has a higher relative proportion of agriculture (6%), suggesting the persistence of primary activities with some structural weight.

Figure 2. Five countries with the highest GDP in the Americas in 2024



Notes: The total figures for the three sectors may not add up to 100% because unallocated consumption has not been captured in the sector-reported data.

Source: Prepared by the author using information from CIA (2024b, 2024a).

4.4. Sectoral distribution in Mexico from 1993 to 2024

From the preceding analysis of the five countries with the highest GDP in the American continent, Mexico stands out for its significant industrial base (31.6%) and strong orientation toward tertiary activities (58.2%). Therefore, due this duality, the analysis now turns to Mexico in greater depth, beginning with its classification as a developing country in Latin America. In 2024, Mexico had a GDP of \$2.883 trillion, ranking 13th worldwide.

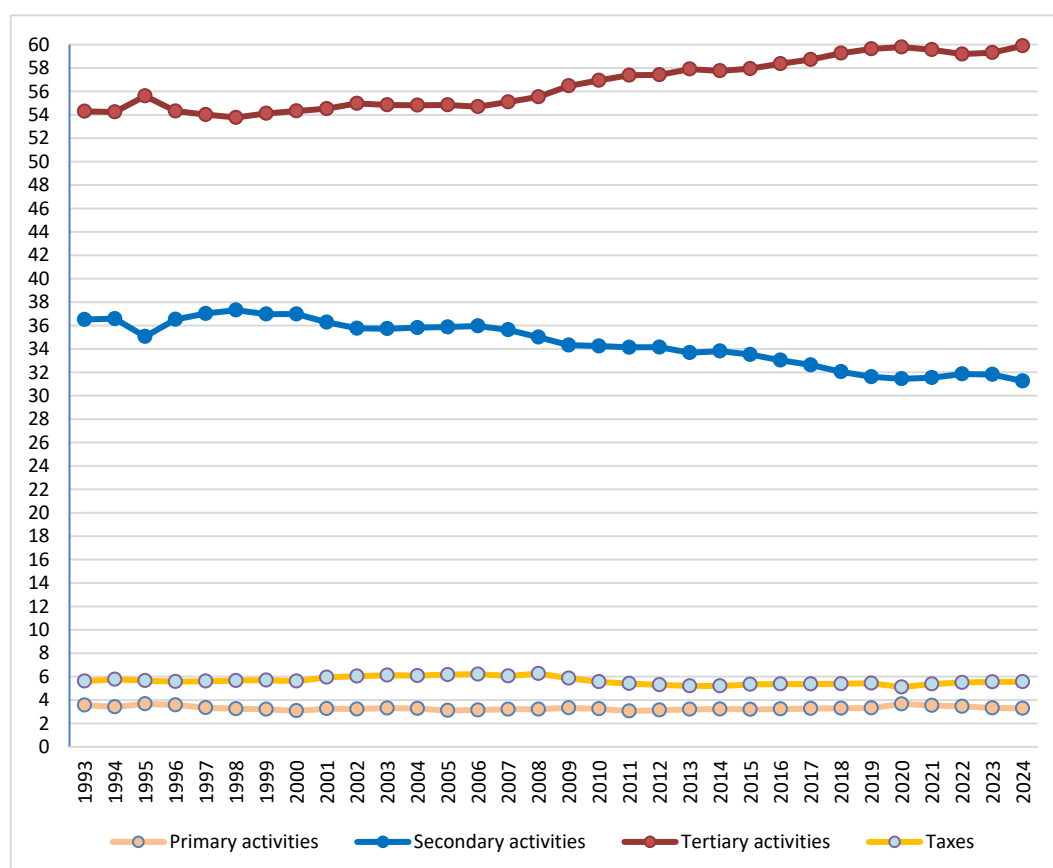
For the past three decades, the Mexican agency that has measured the percentage contribution of different sectors to the economy (GDP) has been the INEGI (2025). It is important to clarify that it does not do so through the World Bank's sectoral classifications (agriculture, industry, and services), but rather through three categories: primary, secondary, and tertiary activities.

Based on the information from INEGI (2025), Mexico's sectoral distribution⁽⁴⁾ from 1993 to 2024 (32 years) (see Figure 3) has maintained a clear predominance of tertiary activities (56.6%), followed by secondary activities (34.5%), while primary activities maintain a marginal weight (3.3%); this is less than taxes on products, which contribute 5.6%. Therefore, the country has had, over the last three decades, an economy primarily oriented towards services, but with a still-relevant industrial base.

A closer examination of the sectoral evolution reveals an upward trend in tertiary activities. After remaining relatively stable at around 54–55% since the 1990s, their share began to increase more markedly in 2009, reaching approximately 57% and approaching 60% by 2024. This pattern indicates a sustained process of tertiarization, further consolidating the services sector as the central component of the economy.

Secondary activities show a downward trend. After reaching levels close to 37% at the end of the 1990s, their share has gradually and almost continuously decreased, settling at around 31% by 2024. The general pattern indicates a relative loss of weight for the industrial sector, consistent with processes of relative deindustrialization and a shift of value added to other sectors.

Figure 3. *Economic sectors (percentage) in Mexico from 1993 to 2024*



Notes: Given the differences in the sectoral classifications of the World Bank and INEGI, a conceptual equivalence is established—with corresponding restrictions—so that agriculture is classified as a primary activity, industry as a secondary activity, and services as a tertiary activity.

Source: Prepared by the author using information from INEGI (2025).

Primary activities remain low and remarkably stable, fluctuating between 3.1% and 3.7%, indicating the absence of significant structural changes in this sector. A slight downward trend is observed from the mid-1990s to the late 2000s, followed by a temporary increase

in 2020, which diminished in subsequent years, returning to its historical average (3.3%). Overall, the data reflects stable performance with marginal variations, confirming the residual nature of the primary sector within the economic structure.

Finally, regarding product taxes, these maintain a relatively stable share, fluctuating between approximately 5.1% and 6.3%. A slight increase is observed during the first half of the 2000s, followed by a gradual reduction from 2009 onwards and a moderate rebound in more recent years, returning to its historical average (5.6%). Overall, their behavior suggests a constant contribution to the total, without substantially altering the general trend of the sectoral structure.

4.5. Sectoral distribution (GDP 2024) by federal entity of Mexico

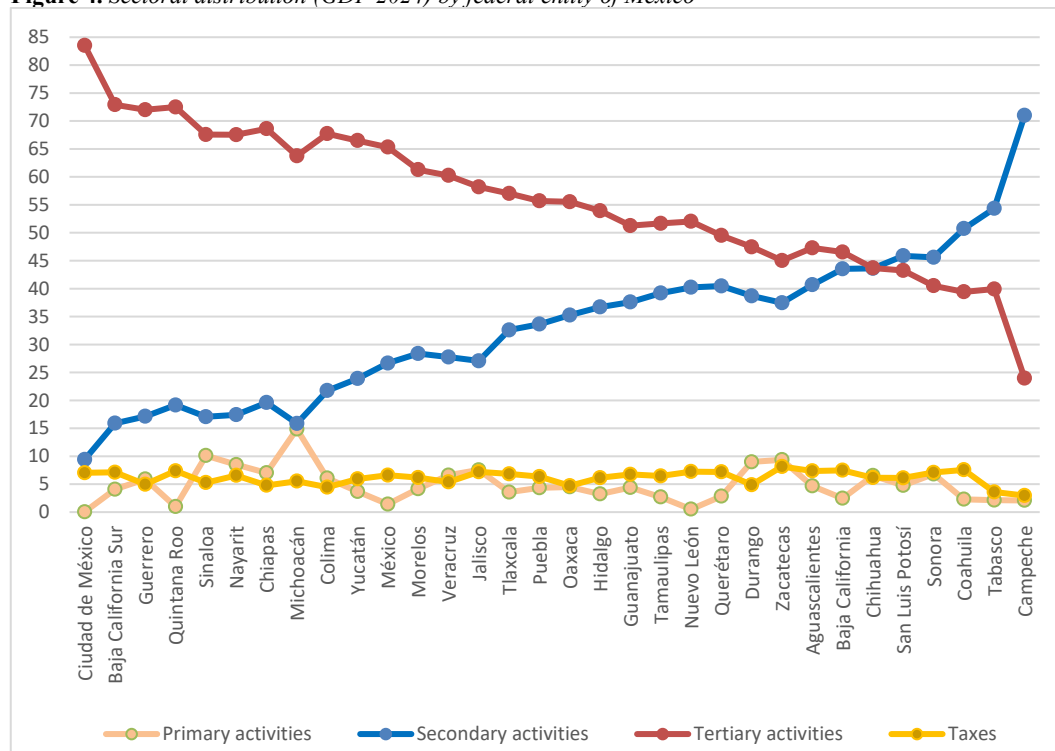
To analyze Mexico's sectoral distribution by federal entity (see Figure 4), the 2024 GDP is used as the base (INEGI, 2024b). Of the 32 federal entities (31 states and the capital, Mexico City) that make up the country, the majority (27 out of 32) have a predominance of the tertiary sector, indicating that the Mexican economy is largely service based at the subnational level. Of the remaining entities, only five have the secondary sector as their main sector; in contrast, none have the primary sector as dominant, which demonstrates the marginal nature of primary activities in the state economic structure.

Of the states with a high concentration of service-based industries, Mexico City stands out as the most extreme case, with 83.52% of its GDP concentrated in tertiary activities, a very small fraction of secondary activities (9.41%), and virtually no participation from primary activities (0.04%). This reflects its role as the main national hub for advanced services and command functions.

Other states with a high degree of specialization in the service sector include Baja California Sur, Quintana Roo, Guerrero, Chiapas, Colima, Sinaloa, and Nayarit, where services account for over 67% (two-thirds of their economies). In several of these cases, the weight of the service sector is strongly associated with tourism and consumer-related services, rather than advanced services, demonstrating different types of tertiarization within the sector itself.

Only five of the 32 states had a predominance of secondary activities: San Luis Potosí, Sonora, Coahuila, Tabasco, and Campeche. In these cases, industry clearly surpasses services; that is, the secondary sector exceeds the tertiary sector. This reached particularly high values in Campeche (70.99%) and Tabasco (54.35%), which is closely related to mining activities (49.01% and 38.93%, respectively) and construction (20.28% in Campeche). In Coahuila, Sonora, and San Luis Potosí, the weight is mainly attributed to manufacturing.

As for primary activities, they account for a small share in practically all states, although participation is relatively higher in states such as Chiapas, Jalisco, Nayarit, Durango, Zacatecas, Sinaloa, and Michoacán (with values ranging from 7.02% to 14.86%). Even so, the primary sector does not constitute the central axis of the state economy, but rather a minor, complementary component.

Figure 4. Sectoral distribution (GDP 2024) by federal entity of Mexico

Source: Prepared by the author using information from the INEGI (2024b).

Overall, Mexico's sectoral structural trend between 1993 and 2024 reveals a sustained process of service sectorization, accompanied by a relative decline in industrial weight and the marginal persistence of the primary sector. This trend is generally observed across states, supporting the 2024 figures, in which over 84% of states exhibit a shift towards service sectorization as their primary sector. However, an analysis of the distribution reveals marked heterogeneity. Consequently, the transition is not uniform but rather demonstrates territorial differentiation that shapes how value is created, specialized, and distributed in the country.

Conclusions

In response to the question of whether a transition from industrial economies to advanced service-based economies has truly occurred, this study assessed the sectoral distribution of 205 countries and territories (CIA, 2024b; 2024a). The results show that over 90% of countries have the service sector as the main component of their GDP. Despite this predominance, it does not imply the disappearance of other sectors within a country or region; rather, one sector acquires a dominant structural weight that decisively influences both individual and collective development. Furthermore, it does not guarantee a homogeneous level of development or a transition to becoming 'advanced countries'.

Indeed, fewer than a quarter are developed countries, almost half are developing countries, and nearly a fifth are least developed countries.

In the specific case of Mexico, sustained growth of the tertiary sector was identified over more than three decades, suggesting a process of relative deindustrialization (INEGI, 2025). Regarding some states with high tertiary specialization, the analysis indicated that this was linked to tourism and consumer-related services rather than advanced services, demonstrating different types of tertiarization within the sector itself. This encourages further research into the effects of different types of tertiarization on the transition and development of economies.

Overall, the data confirms that the shift towards a service-based economy is a dominant global trend; however, they also show that this process is neither uniform nor linear. The high share of services in GDP masks differentiated trajectories, where productive structures maintain hybrid and unequal configurations both between and within countries. In this sense, the expansion of the service sector does not necessarily imply the complete replacement of industrial bases, but rather their reconfiguration within increasingly complex economic frameworks. This makes it imperative to analyze the percentage and internal distribution of these sectors.

The study has two main limitations: First, although the country coverage was broad, it did not include all territories, nor did the figures always cover the entire GDP. Second, it was necessary to establish conceptual equivalencies across sectoral classifications to ensure comparability, since the CIA presents a tripartite typology (agriculture, industry, and services), whereas INEGI categorizes them as primary, secondary, and tertiary activities, as well as taxes. Despite these differences, the general trends observed remain consistent.

For future research, it is recommended to delve deeper into longitudinal analyses that allow for the evaluation of the temporal evolution of this sectoral transition in more countries, as well as to incorporate additional variables such as employment, the degree of productive specialization, and levels of sectoral investment, in order to understand the structural dynamics of economies more accurately. In conclusion, the study contributes to refining the classic narrative of post-industrialization. While tertiarization is predominant globally, it does not constitute an indicator of development, nor does it represent a homogeneous structural transformation. On the contrary, it reflects a diversity of economic trajectories in which industry persists and is articulated in differentiated ways alongside the service sector.

Acknowledgement

I acknowledge the Manuscript Writing Training Team (CEMAI, for its Spanish acronym) of CONACYT (now SECIHTI) for encouraging and supporting me in writing academic texts in English. I also express my sincere gratitude to Dr. Luis Enrique Santiago for fostering my intellectual curiosity about the structural transition from industrial to service-based economies and the implications for urban mobility.

Notes

- (1) Of the total of 205 countries and territories: 48 are developed countries, 100 are developing countries, 41 are least developed countries, and 16 are categorized as special territories. For didactic purposes, from this point forward they will be referred to simply as countries, without diminishing the typology of each other.
- (2) The total figures for the three sectors may not add up to 100% (96.2%) due to unallocated consumption, including financial intermediary services measured indirectly (FISIM) not allocated by industry, and taxes less subsidies on products.
- (3) Except for Japan and South Korea, whose GDP and sectoral data correspond to 2023. Despite Canada's GDP being for 2024, the most recent information available on economic sectors is from 2021.
- (4) Obtained by calculating the average annual percentage for each economic activity and tax from 1993 to 2024.

References

- Bell, D. (2006). *El Advenimiento de La Sociedad Post Industrial* (Cap1). Alianza Editorial.
- Cervelló-Royo, R., Devece, C., Galdón, J.L. and Lull, J.J. (2026). Towards the knowledge society: A definition, measurements, and limits. *Journal of Innovation & Knowledge*, 12, p.100873. <<https://doi.org/10.1016/J.JIK.2025.100873>>
- CIA (2024a). GDP – composition, by sector of origin. The World Factbook. Available at: <<https://www.cia.gov/the-world-factbook/field/gdp-composition-by-sector-of-origin/>>.
- CIA (2024b). Real GDP (purchasing power parity). [online] The World Factbook. Available at: <<https://www.cia.gov/the-world-factbook/field/real-gdp-purchasing-power-parity/>>.
- Clark, C. (1940). The conditions of economic progress. New York: Macmillan Company.
- Fourastié, J. (1949). *Le grand espoir du XXe siècle: progrès technique, progrès économique, progrès social*. Presses universitaires de France.
- Hull, C.H. (1899). *The Economic Writings of Sir William Petty*. Cambridge University Press.
- INEGI (2017). Glosario. Índice de Productividad Laboral y Costo Unitario de la Mano de Obra. Base 2013. Available at: <<https://www.inegi.org.mx/app/glosario/default.html?p=IPL2017>>.
- INEGI (2023). SCIAN México 2023. Sistema de Clasificación Industrial de América del Norte, México SCIAN 2023. Available at: <<https://www.inegi.org.mx/scian/#>>.
- INEGI (2024a). Actividades económicas en México. Available at: <<https://cuentame.inegi.gob.mx/explora/economia/pib/>>.
- INEGI (2024b). Producto Interno Bruto por Entidad Federativa (PIBE). Año base 2018. 2003-2024. [online] Available at: <<https://www.inegi.org.mx/programas/pibent>>.
- INEGI (2025). Producto Interno Bruto Trimestral, de 1993 a 2025. Producto Interno Bruto Trimestral, de 1993 a 2025.
- Krüger, J.J. (2008). Productivity and structural change: A review of the literature. *Journal of Economic Surveys*, 22(2), 330–363. <<https://doi.org/10.1111/j.1467-6419.2007.00539.x>>

- Quiroz Santiago, N.L., Sánchez-Zárate, A. and Santiago, L.E. (2025). Movilidad intraurbana por razones de trabajo en la era digital: Evidencia de la zona metropolitana de Aguascalientes. *Decumanus*, 15(15). <<https://doi.org/10.20983/decumanus.2025.2.8>>
- Schafran, A., McDonald, C., Lopez Morales, E., Akyelken, N. and Acuto, M. (2018). Replacing the services sector and three-sector theory: urbanization and control as economic sectors. *Regional Studies*, 52(12), 1708–1719. <<https://doi.org/10.1080/00343404.2018.1464136>>
- Selstad, T. (1990). The rise of the quaternary sector. The regional dimension of knowledge-based services in Norway, 1970–1985. *Norsk Geografisk Tidsskrift – Norwegian Journal of Geography*, 44(1), 21–37. <<https://doi.org/10.1080/00291959008552242>>
- Sonis, M., Azzoni, C.R. and Hewings, G.J.D. (2008). The Three-sector Growth Hypothesis and the Euler-Malthus Economic growth model: Application to the analysis of GDP dynamics of Brazil, 1985-2004-2020. In: *The Fifth International Conference on Mathematical Modeling and Computer Simulation of Materials Technologies*. 153.

Monetary-fiscal policy regimes and economic growth in Algeria: a Markov-switching analysis (1990Q2–2025Q4)

Moussa AYACHI

University of El Oued, Algeria
moussa-ayachi@univ-eloued.dz

Abdeldjalil CHELLIG

University of El Oued, Algeria
chellig-abdeldjalil@univ-eloued.dz

Abstract. *This paper investigates whether the growth effects of monetary and fiscal policies in Algeria are regime dependent, with particular emphasis on monetary–fiscal policy coordination. Using quarterly data for 1990Q2–2025Q4, we first estimate a baseline linear model and formally test for parameter instability. Structural-break diagnostics (Quandt–Andrews and Chow tests) strongly reject coefficient constancy, indicating that Algeria’s growth process is characterized by discrete shifts that cannot be adequately captured by a single linear specification. Motivated by this evidence, we estimate a two-regime Markov-Switching Dynamic Regression model in which the intercept, policy coefficients, and error variance vary across regimes, while external controls are kept common across regimes. The Markov-switching model provides a substantially better fit than OLS and identifies two persistent regimes that differ in baseline growth and volatility. The central result is that monetary–fiscal policy coordination—measured by the interaction between lagged fiscal and monetary policy indicators—has a positive and statistically significant effect on GDP growth in both regimes, but the effect is markedly stronger the lower-growth, crisis-related regime. Transition probabilities imply long expected durations in both regimes, and smoothed regime probabilities align closely with major shock episodes, notably the 2014 oil-price collapse and the 2020 COVID-19 shock. Overall, the findings highlight the importance of institutionalized policy coordination frameworks, particularly during adverse macroeconomic conditions, while underscoring the growth costs associated with inflationary pressures.*

Keywords: Algeria, economic growth, monetary policy, fiscal policy, policy coordination, Markov-switching, structural breaks.

JEL Classification: E52; E62; E63; O47; C22; O55.

1. Introduction

The relationship between monetary and fiscal policies constitutes one of the most complex and debated issues in macroeconomic research, particularly in resource-rich economies such as Algeria, where understanding the coordinated or uncoordinated impact of these policies on economic stability and growth necessitates a rigorous empirical and theoretical framework (Djaballah, 2024). Algeria is characterized by what is termed the “Algerian Coordination Paradox,” manifested in persistent inflation despite substantial hydrocarbon revenues, and the simultaneous occurrence of fiscal expansion and monetary accommodation during specific periods such as the oil price collapse (2014–2016), the COVID-19 pandemic (2020–2022), and the ongoing digital finance reforms (2024–2026). Each of these episodes has imposed unique challenges requiring nuanced policy responses.

This study aims to examine how event-driven structural shocks influence the effectiveness of monetary–fiscal policy coordination in promoting sustainable economic growth in Algeria (Hanif et al., 2023).

Algeria’s economy is heavily reliant on hydrocarbon exports and has historically faced structural imbalances, with fiscal policy—frequently shaped by oil revenues—playing a pivotal role in determining the budget deficit and the overall economic landscape (Ikram & Mohammed, 2023). In the aftermath of the COVID-19 pandemic, Algeria experienced significant increases in public debt and fiscal pressures, raising concerns about long-term sustainability and implications for macroeconomic stability (Chekouri et al., 2024). These conditions underscore the necessity of analyzing the interaction between fiscal and monetary policies, especially in the presence of exogenous shocks, which can either exacerbate vulnerabilities or present opportunities for structural reforms (Boulila & Benbouziane, 2018).

The “Algerian Coordination Paradox” is defined in this study as the coexistence of large hydrocarbon revenues with fiscal expansion and monetary accommodation, generating persistent inflationary pressures without proportional real economic growth. This structural misalignment is observed across three distinct event regimes: the oil price collapse (2014–2016), the COVID-19 pandemic (2020–2022), and ongoing digital finance reforms (2024–2026), each profoundly shaping policy coordination dynamics (Mahfoudi et al., 2024; Mamman, 2025).

Historically, Algeria’s fiscal policy has been procyclical, expanding during periods of high oil prices and contracting when prices fall, rather than acting as an automatic stabilizer (Ikram & Mohammed, 2023).

This procyclical behavior underscores the critical need for effective coordination between monetary and fiscal policies to cushion the economy against external shocks and promote sustainable development (Mamman, 2025). Furthermore, the COVID-19 pandemic introduced unprecedented challenges, necessitating substantial fiscal stimulus and monetary financing, thereby intensifying coordination challenges and their implications for inflation and growth dynamics (Ikram & Mohammed, 2023).

More recently, institutional reforms associated with the digital transformation of the financial system have introduced an additional dimension to Algeria's macroeconomic policy framework (Jackson, 2024). By modernizing payment systems and strengthening financial infrastructure, these reforms may fundamentally alter the transmission mechanisms and operational frameworks of monetary and fiscal policy. At the same time, the transition toward a more digitalized financial environment may temporarily generate institutional frictions that affect the degree of coordination between economic authorities (Jackson, 2024).

Despite the growing literature on macroeconomic policy in developing and resource-rich economies, most studies tend to analyze monetary and fiscal policies separately. As a result, insufficient attention has been devoted to the joint dynamics of these policies under event-driven structural conditions. In particular, existing research rarely examines how episodic shocks such as commodity price collapses, global crises, or institutional reforms reshape the coordination mechanisms between policy authorities.

This study seeks to determine whether these shocks have exacerbated existing structural imbalances or triggered adaptive policy responses, ultimately influencing the effectiveness of coordinated efforts toward sustainable growth. Expansionary fiscal policies often generate inflationary pressures, necessitating monetary countermeasures, which is especially relevant given Algeria's persistent inflation despite fiscal expansion and monetary accommodation (Bozagliu et al., 2025). This dynamic is further complicated by government programs aimed at economic revitalization, including efforts to control money supply and improve financial stability; however, significant challenges remain in achieving sustainable industrial development (Aid & Benelbar, 2023; Messaoudi et al., 2023).

To address this gap, this study examines the impact of event-driven shocks on the coordination between monetary and fiscal policies in Algeria over the period 1990–2025. The analysis employs a Markov-Switching framework to capture regime-dependent dynamics in monetary–fiscal policy interactions, allowing for endogenous transitions between distinct macroeconomic regimes driven by structural shocks.

Building on the preceding discussion, this study contributes to the literature in three main ways. First, it examines monetary–fiscal policy coordination in Algeria as a resource-dependent economy where hydrocarbon revenues, fiscal pressures, and monetary accommodation jointly shape macroeconomic outcomes. Second, it adopts a nonlinear Markov-Switching framework that allows the effects of fiscal policy, monetary policy, and their interaction on economic growth to vary across unobserved macroeconomic regimes, rather than assuming a stable linear relationship over the whole sample. Third, it distinguishes between relatively stable and crisis-related regimes, thereby assessing whether policy coordination plays a different role during normal macroeconomic conditions compared with periods of stress, such as oil-price shocks and the COVID-19 pandemic. The remainder of the paper is organized as follows. Section II discusses event-driven regime switching in the Algerian context. Section III presents the data and econometric methodology. Section IV reports and discusses the empirical results. The final section concludes and provides policy implications.

This raises a critical theoretical question for the Algerian context:

Does Algeria's persistent inflation reflect a fiscal-led price level determination rather than purely monetary mismanagement?

If oil price shocks reduce expected future surpluses (H1), the model predicts stronger inflationary pressures and weaker coordination effectiveness. Similarly, if extraordinary fiscal expansions during systemic crises—such as pandemics—are not accompanied by credible medium-term consolidation (H2), inflation may emerge as the adjustment variable validating fiscal imbalances.

2. Event-Driven Regime Switching

Macroeconomic policy coordination in Algeria operates within a volatile, resource-dependent environment, where external shocks frequently disrupt fiscal space, monetary policy reaction functions, and the effectiveness of policy coordination. By analyzing these shocks individually, this study aims to better understand the mechanisms through which they induce regime switches between monetary and fiscal dominance.

Oil Price Collapse Regime (2014–2016)

Algeria's economy is heavily reliant on hydrocarbon exports, which historically contribute over 90% of export revenues and approximately 60% of fiscal revenues (International Monetary Fund. Middle East and Central Asia Dept., 2021). A sharp decline in global oil prices between 2014 and 2016 led to a contraction in fiscal revenues of approximately 30–35%, widening budget deficits and reducing fiscal space for countercyclical spending. This environment created conditions conducive to fiscal dominance, as the government faced political and social pressures to maintain public spending, often financed through central bank borrowing or debt monetization. (Elkhdari et al., 2018)

From a theoretical perspective, the oil shock represents a regime-dependent exogenous variable that alters the effectiveness of monetary policy instruments. The fiscal constraint imposed by declining revenues modifies the central bank's reaction function, potentially forcing it to accommodate inflationary pressures to prevent sovereign default risk. Empirically, inflation in Algeria remained elevated (approximately 6–8%) during this period despite tight monetary measures, indicating that fiscal shocks dominated price level determination, consistent with the Fiscal Theory of the Price Level (FTPL) (Ikram & Mohammed, 2023; Mehibel & Belarbi, 2018)

COVID-19 Pandemic Regime (2020–2022)

The COVID-19 pandemic introduced unprecedented macroeconomic and fiscal challenges globally and in Algeria. In 2020, Algeria implemented massive fiscal stimulus programs, including liquidity injections and subsidies, which temporarily increased the fiscal deficit by more than 4% of GDP (International Monetary Fund. Middle East and Central Asia Dept., 2021). Simultaneously, the pandemic triggered disruptions in global trade, oil demand, and domestic economic activity, amplifying uncertainty in both monetary and fiscal policy transmission mechanisms (Ikram & Mohammed, 2023).

In this regime, the policy environment shifted dynamically between active fiscal interventions and constrained monetary policy responses (Ferdj, 2024). The central bank's ability to stabilize inflation and the real economy was partially constrained by the necessity to accommodate fiscal stimulus without jeopardizing public debt sustainability. This scenario illustrates a temporary regime-switching phenomenon, where fiscal actions dominated short-term economic outcomes, while the central bank retained partial autonomy to manage inflation expectations, highlighting the asymmetric interaction between the two authorities. (Campos et al., 2025)

Digital Finance Reform Regime (2024–2026)

Ongoing digital finance reforms in Algeria aim to modernize payment systems, enhance financial inclusion, and improve liquidity management (Mahfoudi et al., 2024). This regime represents a structural institutional shock that affects both fiscal and monetary policy implementation. Digitalization expands the central bank's operational capacity while also increasing transparency and accountability in fiscal spending, potentially reducing unmonitored fiscal dominance. (Dionysopoulos et al., 2024; Genc & Takagi, 2024)

From a theoretical standpoint, these reforms alter the policy transmission channels, creating opportunities for improved coordination between fiscal and monetary authorities. However, the transition phase introduces temporary frictions, where legacy fiscal practices may still clash with modern monetary management tools. This justifies the adoption of a Markov-Switching framework, allowing the model to capture nonlinear regime-dependent effects on inflation, money supply, and growth dynamics.

Building on the historical and institutional analysis presented above, this study develops a regime-dependent monetary–fiscal coordination framework grounded in a Markov-Switching paradigm. Rather than modeling external shocks—such as oil price collapses, global crises, or institutional reforms—through deterministic and exogenously imposed dummy variables, which implicitly assume that structural breaks are perfectly known and discrete, this framework treats such events as underlying structural catalysts that induce endogenous transitions between unobserved macroeconomic regimes.

Accordingly, the Algerian economy is assumed to evolve under a set of latent regimes S_t , which capture distinct patterns of policy interaction (e.g., monetary dominance, fiscal dominance, or hybrid coordination regimes). These regimes are not directly observable but are inferred probabilistically from the data, reflecting the inherently nonlinear and time-varying nature of macroeconomic dynamics.

Formally, the theoretical specification can be represented as a regime-dependent interaction model:

$$Y_t = \alpha_{S_t} + \beta_{1,S_t} FB_t + \beta_{2,S_t} MS_t + \beta_{3,S_t} (FB_t \times MS_t) + \gamma X_t + \varepsilon_t, \varepsilon_t \sim N(0, \sigma_{S_t}^2)$$

where:

Y_t denotes the macroeconomic outcome variable (e.g., economic growth or inflation),
 FB_t represents the fiscal policy stance (fiscal balance),

MS_t denotes the monetary policy stance (money supply),
 $FB_t \times MS_t$ captures the degree of policy coordination,
 X_t is a vector of control variables (e.g., oil prices, exchange rate, inflation),
 $S_t \in \{1, 2, \dots, K\}$ is an unobserved regime variable indicating the prevailing macroeconomic regime at time t ,
 $\alpha_{S_t}, \beta_{1,S_t}, \beta_{2,S_t}, \beta_{3,S_t}$, and $\sigma_{S_t}^2$ are regime-dependent parameters.

The latent regime variable S_t evolves according to a first-order Markov process:

$$P(S_t = j \mid S_{t-1} = i) = p_{ij}, \sum_{j=1}^K p_{ij} = 1$$

This specification implies that the probability of transitioning between regimes depends solely on the previous regime, thereby capturing the persistence and stochastic nature of macroeconomic regimes.

From a theoretical perspective, this regime-dependent structure reflects the notion that policy effectiveness, dominance structures, and coordination mechanisms are inherently state-contingent. For instance, adverse oil price shocks may compress fiscal revenues and induce a transition toward a fiscal dominance regime, wherein monetary policy becomes accommodative. Conversely, institutional reforms—such as financial digitalization—may enhance policy transmission and facilitate a transition toward a more stable and coordinated regime.

Importantly, the Markov-Switching framework allows these regime changes to be identified endogenously, without imposing prior assumptions regarding their timing or duration. This provides a significant methodological advantage over conventional approaches based on deterministic structural breaks, as it enables the data to reveal the underlying regime dynamics.

Consequently, this specification offers a robust and flexible empirical representation of macroeconomic behavior in Algeria, allowing for the identification of whether economic outcomes are driven by monetary dominance, fiscal dominance, or transitional hybrid regimes, in line with the theoretical predictions of nonlinear policy interaction models.

3. Materials and Methods

3.1. Data and Variables Description

This study employs a quarterly multivariate dataset covering the period 1990Q1–2025Q4 to examine the regime-dependent dynamics of monetary–fiscal policy coordination in Algeria. Given the limited availability of official quarterly fiscal data in the early part of the sample, annual series are temporally disaggregated using the Denton–Cholette method and cubic spline interpolation. These procedures preserve the long-run properties of the original data while allowing for the construction of higher-frequency observations suitable for capturing nonlinear dynamics and potential regime shifts.

All variables are transformed where necessary (growth rates and log differences) to ensure stationarity. The stationarity properties of the series are confirmed using Augmented Dickey–Fuller (ADF) and Phillips–Perron (PP) tests.

The dataset includes the following key variables:

The dependent variable is real GDP growth ($GDPG_t$), which captures short-term fluctuations in economic activity and serves as the main indicator of macroeconomic performance across regimes.

Fiscal policy (FB_t) is measured by the overall fiscal balance as a percentage of GDP. Positive values indicate fiscal surpluses (contractionary stance), while negative values reflect fiscal deficits (expansionary stance). In the Algerian context, this variable is closely linked to hydrocarbon revenues and external shocks.

Monetary policy (MS_t) is proxied by the growth rate of broad money (M2). Given the limited role of interest rates in the Algerian monetary transmission mechanism, money supply growth represents the primary operational instrument of monetary policy.

The policy interaction term ($FB_t \times MS_t$) captures the joint effect of fiscal and monetary policies. A positive coefficient is interpreted as policy complementarity (coordination), while a negative coefficient suggests policy conflict. Within the Markov-Switching framework, this interaction is allowed to vary across unobserved regimes, capturing potential regime dependence in policy coordination effects.

Finally, a set of control variables (X_t) is included to account for external and structural influences. These include global oil prices (OIL_t), the nominal exchange rate (EXR_t), and the inflation rate (INF_t), which jointly capture external shocks, competitiveness effects, and macroeconomic stability conditions.

3.2. Econometric Specification: Markov-Switching Framework

To capture nonlinear and regime-dependent interactions between monetary and fiscal policies in Algeria, this study employs a Markov-Switching Dynamic Regression (MS-DR) model. Unlike linear specifications, this framework allows model parameters to vary across unobserved regimes governed by a latent regime variable S_t .

Let $S_t \in \{1, 2, \dots, K\}$ denote the unobserved regime at time t . To mitigate potential endogeneity arising from the simultaneous interaction between economic growth and policy decisions, fiscal and monetary variables are introduced with a one-period lag.

The regime-dependent growth equation is specified as:

$$GDPG_t = \alpha_{S_t} + \beta_{1,S_t} FB_{t-1} + \beta_{2,S_t} MS_{t-1} + \beta_{3,S_t} (FB_{t-1} \times MS_{t-1}) + \gamma X_t + \varepsilon_t$$

where:

α_{S_t} : regime-specific intercept

$\beta_{1,S_t}, \beta_{2,S_t}, \beta_{3,S_t}$: regime-dependent coefficients for fiscal policy, monetary policy, and their interaction

X_t : vector of control variables (oil prices, exchange rate, inflation)

ε_t : error term

To account for regime-dependent volatility, the disturbance term is specified as:

$$\varepsilon_t \sim N(0, \sigma_{S_t}^2)$$

This specification allows both the conditional mean and variance of economic growth to vary across regimes, thereby capturing differences in macroeconomic stability and shock absorption capacity between periods of stability and crisis.

3.3. Markov Transition Matrix (Regime-Switching Dynamics)

The evolution of the unobserved regime variable S_t is assumed to follow a first-order Markov process, implying that the probability of being in a given regime at time t depends only on the state at time $t - 1$, and is independent of the full historical path.

Assuming a two-regime structure (Regime 1; Regime 2), the transition dynamics are characterized by the Markov transition matrix:

$$P = \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix}$$

where the transition probabilities are defined as:

$$p_{ij} = P(S_t = j \mid S_{t-1} = i)$$

Specifically:

p_{11} : probability of remaining in Regime 1

$p_{12} = 1 - p_{11}$: probability of transitioning from Regime 1 to Regime 2

p_{22} : probability of remaining in Regime 2

$p_{21} = 1 - p_{22}$: probability of transitioning from Regime 2 back to Regime 1

High values of the diagonal elements (p_{11} and p_{22}) indicate strong regime persistence, implying that once the economy enters a given macroeconomic state, it is likely to remain in that state for a prolonged period.

In the case of Algeria, these transition probabilities provide a probabilistic representation of regime stability and exposure to external shocks, such as oil price fluctuations, global financial disruptions, or pandemic-related shocks, which may induce transitions between Regime 1 and Regime 2.

3.4. Estimation Strategy: Maximum Likelihood and Filtering Approach

The empirical estimation of the Markov-Switching model is based on Maximum Likelihood Estimation (MLE), which enables the joint estimation of regime-dependent parameters and the probabilities of the unobserved state variable S_t . Since the regime is not directly observable, the likelihood function is constructed by integrating over all possible regime paths.

To infer the latent regimes, the study employs the Hamilton (1989) filtering algorithm, a recursive procedure that updates the probability of each regime conditional on information available up to time t . This filtering approach allows the model to track time-varying regime probabilities in a consistent and data-driven manner.

The log-likelihood function is maximized using numerical optimization techniques, specifically the Broyden–Fletcher–Goldfarb–Shanno (BFGS) algorithm. To improve estimation robustness and mitigate the risk of convergence to local maxima, a random search over multiple sets of initial parameter values is implemented. Initial regime probabilities are obtained from the ergodic (long-run) distribution of the Markov chain, ensuring consistency between initial conditions and the estimated transition dynamics.

This likelihood-based estimation strategy allows both parameter heterogeneity and regime transitions to be fully determined endogenously from the data. Consequently, the framework is able to capture genuine structural changes in the Algerian economy—such as those associated with the 2014 oil price collapse and the COVID-19 pandemic—without relying on exogenously imposed dummy variables.

4. Results and Discussion

4.1 Baseline Linear Estimates and the Rationale for Non-Linearity

A baseline OLS regression was first estimated as a benchmark under the maintained assumption of stable parameters over the full sample.

Table 1: *Baseline OLS Estimation Results (1990Q1–2025Q4)*

Variable	Coefficient	Std. Error	T-Statistic	Prob
C	3.007715	1.08000	2.784921	0.0061
FB (-1)	-0.504397	0.167251	-3.015810	0.0031
MS (-1)	-0.097299	0.019189	-5.070426	0.0000
FB_MS_LAG1	0.008164	0.002731	2.986752	0.0033
OIL	0.130059	0.043545	2.986752	0.0033
EXR	0.048681	0.008664	5.619028	0.0000
INF	-0.040662	0.025027	-1.624743	0.1065
DW	0.161946			

Source:

While several coefficients appear statistically significant, the linear specification is empirically fragile. In particular, the Durbin–Watson statistic is extremely low ($DW \approx 0.16$), indicating strong residual persistence and suggesting that a constant-parameter model is unable to capture relevant shifts in the underlying data-generating process.

To distinguish between simple serial correlation and genuine parameter instability, formal breakpoint tests were conducted.

Table 2: *Structural Break Tests for the Baseline OLS Equation (Quandt–Andrews and Chow Tests)*

Quandt–Andrews unknown breakpoint test		
Statistic	Value	Prob
Maximum LR F-Statistic (2013Q1)	39.92077	0.0000
Maximum Wald F-Statistic (2013Q1)	279.4454	0.0000
Chow breakpoint test		
F-Statistic	Prob	
5.042713	0.0000	
38.49927	0.0000	
38.188087	0.0000	

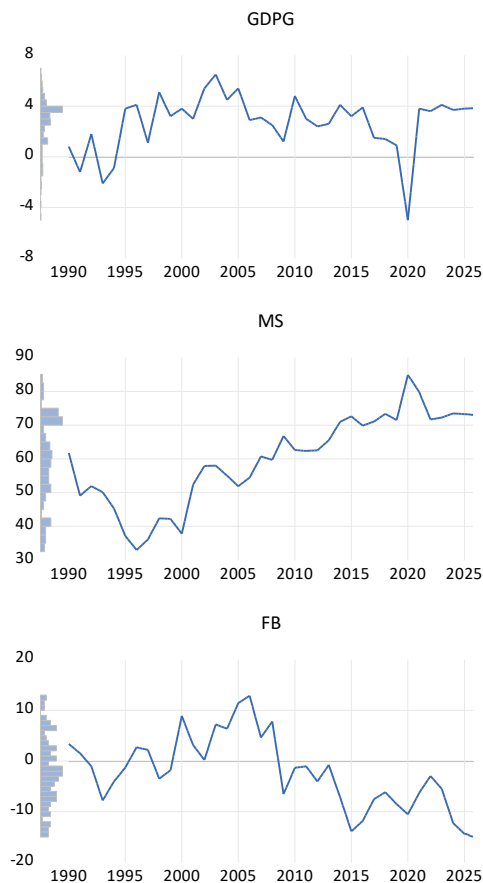
Source:

reports the Quandt–Andrews unknown breakpoint test, which strongly rejects the null hypothesis of no breaks within the trimmed sample (all test statistics significant at the 1% level), with the maximum statistic indicating a likely break around 2013Q1.

In addition, Chow breakpoint tests at economically meaningful dates 2014Q3/2014Q4 (the onset of the oil-price collapse) and 2020Q1 (the COVID-19 shock) also reject parameter constancy at conventional significance levels (p -values < 0.01).

Overall, the baseline diagnostics and breakpoint evidence indicate that Algeria's growth process is characterized by discrete shifts and regime-dependent parameters. This provides direct empirical motivation for adopting a Markov-switching framework in the subsequent analysis.

Figure 1: Time-Series Plots of GDP Growth and Policy Variables (GDPG, FB, MS)



Before exploring the state-dependent dynamics of monetary and fiscal policy, an initial linear Ordinary Least Squares (OLS) model was estimated to serve as a baseline reference. The linear specification assumes that the macroeconomic relationships and policy transmission mechanisms in Algeria remained constant throughout the entire 1990–2025 sample period, effectively ignoring the potential impact of structural shocks.

4.2. Markov-Switching Estimation Results and Regime Identification

A two-regime Markov-Switching Regression model was estimated to capture potential nonlinear dynamics in Algeria's output growth. The model was estimated using maximum likelihood techniques with the BFGS/Marquardt optimisation algorithm and converged successfully after 56 iterations, indicating a stable and well-behaved likelihood function. The specification allows for regime-dependent intercepts and variance structures, while maintaining common slope coefficients for macroeconomic variables across regimes.

Table 3: *Markov-Switching Dynamic Regression (MS-DR) Estimation Results (Two-Regime Model), 1990Q2–2025Q4*

Panel A. Regime-dependent coefficients

Variable	Regime 1 Coef.	Std. Err.	z-Stat	p-value	Regime 2 Coef.	Std. Err.	z-Stat	p-value
C	8.047184	1.416757	5.680003	0.0000	4.717387	0.758026	6.223255	0.0000
FB (-1)	-0.51029	0.220508	-2.31416	0.0207	-0.84694	0.133869	-6.32666	0.0000
MS (-1)	-0.11040	0.015976	-6.91079	0.0000	-0.09083	0.019975	-4.54767	0.0000
FB MS LAG1	0.007977	0.003230	2.469420	0.0135	0.015217	0.002519	6.039994	0.0000
LOG(SIGMA)	-0.37922	0.103078	-3.67903	0.0002	-0.02252	0.096513	-0.23343	0.8154

Panel B. Common (non-switching) regressors

Variable	Coefficient	Std. Error	z-Statistic	p-value
OIL	0.063536	0.045438	1.398309	0.1620
EXR	0.034363	0.005010	6.858371	0.0000
INF	-0.127204	0.018915	-6.725037	0.0000

Panel C. Model fit and estimation information

Item	Value
Included observations	143
Number of regimes	2
Convergence	Achieved after 56 iterations
Log-likelihood	-202.3562
AIC	3.039946
BIC (Schwarz)	3.350734
Hannan–Quinn	3.166236
S.E. of regression	1.056574
Durbin–Watson	0.442254

The estimation results reported in Table 3 confirm the presence of two statistically distinct regimes with differences in intercepts, dynamic coefficients, and variance structures. Most regime-dependent parameters are statistically significant, while the common regressors display heterogeneous levels of significance. Overall, the results provide strong evidence of regime-dependent dynamics in Algeria's economic growth process.

4.3. Regime Characterization and Economic Interpretation

This section provides an economic interpretation of the two latent regimes identified by the MS-DR model and discusses the role of the common control variables. While Section 4.2 established the presence of statistically distinct regimes, the present section characterizes their macroeconomic features and examines how key fundamentals shape growth across regimes.

Regime 1:

Regime 1 is characterized by a high and statistically significant intercept ($C = 8.047$, $p < 0.01$), indicating strong underlying growth momentum when the economy operates in this regime. In addition, the estimated log-variance term is negative and statistically significant ($\text{LOG}(\text{SIGMA}) = -0.379$, $p < 0.01$), suggesting relatively low output volatility and a more stable macroeconomic environment.

Economically, this regime is consistent with periods in which macroeconomic conditions are relatively predictable, often associated with favorable external conditions and improved fiscal space in a hydrocarbon-dependent economy.

Within this regime, both lagged fiscal and monetary policy variables are statistically significant, and the interaction term remains positive and significant. This suggests that policy actions and their interaction contribute to growth dynamics under relatively stable macroeconomic conditions.

Regime 2:

Regime 2 exhibits a substantially lower (though still statistically significant) intercept ($C = 4.717$, $p < 0.01$), indicating weaker baseline growth performance relative to Regime 1. The estimated log-variance term is close to zero and statistically insignificant ($\text{LOG}(\text{SIGMA}) = -0.023$, $p = 0.815$), which is consistent with a more uncertain and less stable macroeconomic environment.

From an economic perspective, this regime corresponds to periods of macroeconomic stress, where adverse shocks and structural constraints—such as fiscal pressures or unfavorable external conditions—limit growth performance.

Importantly, fiscal and monetary policy coefficients remain statistically significant in this regime, indicating that policy transmission mechanisms persist even under adverse conditions, although they operate within a context of weaker fundamentals and higher uncertainty.

Common Control Variables (Non-Switching Regressors)

The control variables are specified as common across regimes, implying that their marginal effects are assumed to be stable across different macroeconomic regimes.

Oil prices (OIL): The estimated coefficient is positive but statistically insignificant at conventional levels (0.0635 ; $p = 0.162$). This suggests that oil-price effects may be transmitted indirectly through policy responses or regime shifts rather than through a stable within-regime marginal effect.

Exchange rate (EXR): The exchange-rate coefficient is positive and highly significant (0.0344 ; $p < 0.01$), indicating a robust association between exchange-rate movements and output growth. This relationship may reflect competitiveness effects as well as broader macroeconomic adjustment mechanisms.

Inflation (INF): Inflation exerts a strongly negative and statistically significant effect on growth (-0.1272 ; $p < 0.01$), consistent with the view that persistent inflation undermines

real economic performance through its impact on purchasing power, investment, and macroeconomic stability.

4.5. Transition Dynamics and Historical Shock Alignment

This section examines the persistence of the estimated regimes and the likelihood of switching between them. The Markov transition probabilities summarize the degree of regime persistence, while the implied expected duration measures how long the economy typically remains in each regime once entered. These results provide a bridge between the latent regimes identified by the model and Algeria's macroeconomic history.

Table 4: *Transition Probabilities and Expected Duration*

Panel A. Transition Probability Matrix

From / To	Regime 1	Regime 2
Regime 1	0.932	0.068
Regime 2	0.078	0.922

Panel B. Expected Duration of Regimes

Regime	Expected Duration (Quarters)	Approx. Years
Regime 1	14.76	3.7
Regime 2	12.85	3.2

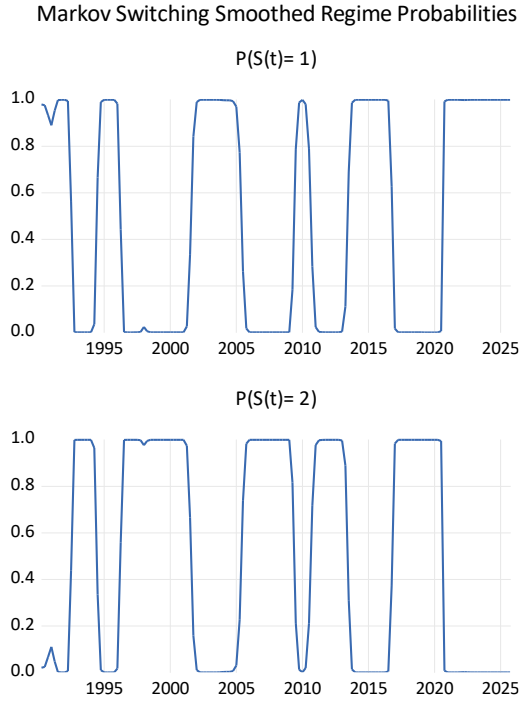
The estimated transition matrix indicates a high degree of persistence in both regimes. The probability of remaining in Regime 1 conditional on being in that Regime is high ($P_{11} = 0.932$), implying that periods of relative macroeconomic stability tend to be sustained over time. The corresponding expected duration is approximately $D_1 = 1/(1 - P_{11}) \approx 14.76$ quarters (about 3.7 years).

Similarly, the probability of remaining in Regime 2 is also high ($P_{22} = 0.922$), with an expected duration of approximately $D_2 \approx 12.85$ quarters (about 3.2 years). These results suggest that once the economy enters a stress regime, it tends to remain there for a prolonged period, consistent with the persistence of macroeconomic imbalances in resource-dependent economies.

The off-diagonal elements of the transition matrix capture switching behavior between regimes. The probability of moving from Regime 1 to Regime 2 is relatively low ($P_{12} = 0.068$), while the probability of transitioning from Regime 2 back to Regime 1 is slightly higher ($P_{21} = 0.078$). Although both probabilities remain modest, this pattern suggests that recoveries are feasible but tend to occur gradually rather than abruptly.

To relate these latent dynamics to observable history, the smoothed regime probabilities provide a time-varying measure of the likelihood that the economy is in each regime.

Figure 2: *Smoothed regime probabilities of Regime 1 and Regime 2 in the Markov-Switching Model (Probability of Regime 1 and Regime 2)*



Periods in which the probability of Regime 2 approaches one can be interpreted as episodes where the economy is more likely to be in a stress regime. These intervals broadly coincide with major macroeconomic disturbances, including the decline in oil prices around 2014 and the global disruptions associated with the COVID-19 pandemic.

Overall, the transition analysis confirms that Algeria alternates between two persistent regimes, with stress episodes exhibiting substantial inertia once triggered. Combined with the regime-dependent coordination effects documented in Section 4.4, these findings suggest that policy coordination may be particularly important during prolonged stress periods, when the economy's capacity to self-stabilize is relatively limited.

Conclusion

This paper examined the growth effects of monetary and fiscal policies in Algeria with a particular focus on whether policy coordination is regime-dependent. Using quarterly data over 1990Q2–2025Q4, baseline OLS estimates and structural-break diagnostics provided strong evidence against parameter stability. The Quandt–Andrews and Chow tests reject the null of no breaks, indicating that a constant-coefficient linear specification is not adequate for capturing Algeria's evolving macroeconomic environment.

Motivated by this evidence, the analysis employed a two-regime Markov-Switching Dynamic Regression (MS-DR) model. The Markov specification fits the data substantially

better than the linear benchmark and identifies two persistent regimes that differ in baseline growth and volatility. Most importantly, the results show that the growth impact of monetary–fiscal coordination is positive and statistically significant in both regimes, but markedly stronger in the crisis regime. This confirms that coordination is not merely a “normal-times” optimization tool; rather, it becomes particularly consequential when the economy is subject to adverse shocks and heightened uncertainty.

The transition probabilities indicate that both regimes are highly persistent, and that crisis episodes, once triggered, tend to last multiple quarters. The smoothed regime probabilities align closely with historically salient shocks—most notably the 2014 oil-price collapse and the 2020 COVID-19 shock—supporting the interpretation that the estimated regimes capture meaningful shifts in Algeria’s macroeconomic conditions.

From a policy perspective, these findings suggest that strengthening institutional mechanisms for monetary–fiscal coordination can yield growth benefits, especially during periods of macroeconomic stress when policy inconsistency is more costly. At the same time, the consistently negative effect of inflation highlights the importance of anchoring price stability objectives within any coordination framework to avoid growth-eroding inflationary outcomes.

This study is not without limitations. The regime-switching framework remains a reduced-form representation of complex policy interactions, and alternative measures of fiscal stance, monetary conditions, and external vulnerability could further refine the analysis. Future research could extend the model by allowing additional regimes, incorporating time-varying transition probabilities linked to oil-market indicators, or exploring nonlinear effects on other macroeconomic outcomes such as inflation and unemployment.

References

- Aid, L., and Benelbar, M. (2023). Study of the Standard Relationship between the Money Supply and the Exchange Rate in Algeria during the Period (1990/2020). *Financial Markets, Institutions and Risks*, 7(2), 56–71. <[https://doi.org/10.21272/fmir.7\(2\).56-71.2023](https://doi.org/10.21272/fmir.7(2).56-71.2023)>
- Boulila, H., and Benbouziane, M. (2018). Austerity in Time of Crisis: A Solution or a Dangerous Idea? Evidence from Algeria. *ETIKONOMI*, 17(1), 11–24. <<https://doi.org/10.15408/etk.v17i1.6799>>
- Bozagiu, A.-M., Mihai, D.-G., and Toader, C.-I. (2025). *Fiscal and monetary policy interactions: impacts and economic implications*.
- Campos, M.M., Cardoso-Costa, J., Gomes, S., and Jacquinot, P. (2025). Monetary and Fiscal Policy Interactions in the Aftermath of an Inflationary Shock. *SSRN Electronic Journal*. <<https://doi.org/10.2139/ssrn.5715583>>
- Chekouri, S.M., Chibi, A., and Benbouziane, M. (2024). Public debt dynamics and fiscal sustainability in selected North African countries: New evidence from recurrent explosive behavior tests and quantile unit root analysis. *Economic Change and Restructuring*, 57(2), 38. <<https://doi.org/10.1007/s10644-024-09625-w>>

- Dionysopoulos, L., Marra, M., and Urquhart, A. (2024). Central bank digital currencies: A critical review. *International Review of Financial Analysis*, 91, 103031. <<https://doi.org/10.1016/j.irfa.2023.103031>>
- Djaballah, M. (2024). Fiscal Policy and Economic Growth a Cointegration Approach Based on Structural Breaks: Evidence from Algeria. *Applied Economics and Finance*, 11(1), 45. <<https://doi.org/10.11114/ae.v11i1.6775>>
- Elkhdari, M., Souissi, M., and Jewell, A. (2018). Empirical Estimation of Fiscal Multipliers in MENA Oil-Exporting Countries with an Application to Algeria. *IMF Working Papers*, 18(124), 1. <<https://doi.org/10.5089/9781484357972.001>>
- Ferdj, Y. (2024). The impact of the COVID-19 pandemic on regional economic integration in the Maghreb region – an Algerian perspective. *Economic and Regional Studies / Studia Ekonomiczne i Regionalne*, 17(2), 290–313. <<https://doi.org/10.2478/ers-2024-0016>>
- Genc, H. O., and Takagi, S. (2024). A literature review on the design and implementation of central bank digital currencies. *International Journal of Economic Policy Studies*, 18(1), 197–225. <<https://doi.org/10.1007/s42495-023-00125-9>>
- Hanif, A., Salah Uddin, M., Bakirtas, T., and Kader, S. A. (2023). Military Expenditure and Unemployment Nexus in Bangladesh. *Revista Finanzas y Política Económica*, 15(1), 157–184. <<https://doi.org/10.14718/revfinanzpolitecon.v15.n1.2023.7>>
- Ikram, B., and Mohammed, K. S. (2023). The Effects of Fiscal Dominance on Monetary Policies in Algeria Amidst COVID-19. *Zagreb International Review of Economics and Business*, 26(2), 77–96. <<https://doi.org/10.2478/zireb-2023-0015>>
- International Monetary Fund. Middle East and Central Asia Dept. (2021). Algeria: 2021 Article IV Consultation-Press Release; and Staff Report; and Statement by the Executive Director for Algeria. *IMF Staff Country Reports*, 2021(253), 1. <<https://doi.org/10.5089/9781616357948.002>>
- Jackson, E. A. (2024). Economics of Fiscal Dominance and Ramifications for the Discharge of Effective Monetary Policy Transmission. *Theoretical and Practical Research in Economic Fields*, 15(1), 31. <[https://doi.org/10.14505/tpref.v15.1\(29\).03](https://doi.org/10.14505/tpref.v15.1(29).03)>
- Mahfoudi, F., Riache, S., and Louail, B. (2024). The effect of natural resources and foreign direct investment on economic growth. *Economics and Environment*, 88(1), 731. <<https://doi.org/10.34659/eis.2024.88.1.731>>
- Mamman, S. O. (2025). Role of Fiscal-Monetary Policy Coordination in Managing Energy Price Shocks in Oil-Exporting Small-Open Economy. *Energy RESEARCH LETTERS*, 7(Early View). <<https://doi.org/10.46557/001c.138290>>
- Mehibel, S., and Belarbi, Y. (2018). Inflation Dynamics in Algeria. *Journal of Economics and Behavioral Studies*, 9(6), 174–187. <<https://doi.org/10.22610/jeb.v9i6.2014>>
- Messaoudi, A., Derbal, F. Z., Hasnaoui, M., and Belhamidi, H. (2023). How far does the central bank influence its economies? The example of Algeria. *SocioEconomic Challenges*, 7(2), 45–53. <[https://doi.org/10.21272/sec.7\(2\).45-53.2023](https://doi.org/10.21272/sec.7(2).45-53.2023)>

Does foreign aid reach the poorest? A multidimensional poverty analysis for Pakistan

Noel ALTER

Higher Education Department, Government of Punjab, Pakistan
noelalter@yahoo.com

Arslan KHALID

Higher Education Department, Government of Punjab, Pakistan
arslan.khalid999@gmail.com

Abstract. *This study examines whether foreign aid effectively reaches the poorest segments of society by analyzing its impact across multiple dimensions of poverty in Pakistan. Moving beyond conventional approaches that rely solely on the poverty headcount ratio, this research employs the Foster-Greer-Thorbecke class of poverty measures, including poverty incidence, poverty gap, and poverty severity, to capture distributional dynamics. Using panel data spanning 1995-2025 and applying a dynamic System Generalized Method of Moments estimator, the study addresses issues of endogeneity, persistence, and unobserved heterogeneity. The results reveal that foreign aid significantly reduces poverty incidence but exhibits diminishing effectiveness when addressing poverty depth and severity. This suggests that aid disproportionately benefits individuals closer to the poverty line, while the poorest remain relatively underserved. The findings contribute to the literature by highlighting the distributional limitations of foreign aid and underscore the need for more targeted and inclusive policy interventions.*

Keywords: multidimensional, poverty, severity, inequality, GMM.

JEL Classification: F35, I32.

1. Introduction

Poverty reduction remains at the core of global development agendas, as reflected in the Millennium Development Goals (MDGs) and the Sustainable Development Goals (SDGs) (World Bank, 2023; Haughton & Khandker, 2009). Over the past several decades, foreign aid has been one of the primary instruments used by the international community to combat poverty in developing countries (Burnside & Dollar, 2000; Collier & Dollar, 2002). Despite substantial financial commitments, however, the effectiveness of foreign aid in achieving meaningful and inclusive poverty reduction remains a subject of ongoing debate (Dollar & Kraay, 2004).

A major limitation in this debate lies in how poverty is measured. Most empirical studies rely on the poverty headcount ratio, which measures the proportion of individuals living below a specified poverty line. While this measure provides a useful snapshot of poverty incidence, it fails to capture important aspects such as the depth of poverty and inequality among the poor. As a result, it may lead to an overly optimistic assessment of development outcomes by ignoring the conditions of those who remain far below the poverty line.

The introduction of the Foster-Greer-Thorbecke (FGT) class of poverty measures has significantly improved the analysis of poverty by allowing researchers to examine its multiple dimensions (Easterly, 2006). The poverty gap index measures the average shortfall of income from the poverty line, while the severity index captures inequality among the poor. Together, these measures provide a more comprehensive understanding of poverty dynamics and allow for a more nuanced evaluation of policy interventions (Foster et al. 1984).

This study builds on this framework to examine whether foreign aid effectively reaches the poorest segments of society in Pakistan. The central argument is that while foreign aid may reduce overall poverty levels, its impact may not be evenly distributed. In particular, aid may be more effective in lifting individuals who are close to the poverty line, while having limited impact on those who are deeply impoverished.

Pakistan provides a compelling case for this analysis due to its long history of foreign aid inflows and persistent poverty challenges. Despite periods of economic growth and significant aid receipts, the country continues to face high levels of poverty, inequality, and regional disparities (Alvi & Senbeta, 2012). These conditions raise important questions about the distributional effectiveness of foreign aid and the extent to which it benefits the most vulnerable populations.

By employing a dynamic panel data approach and using multiple poverty measures, this study aims to provide a comprehensive assessment of the impact of foreign aid on poverty in Pakistan. The findings contribute to the literature by highlighting the importance of considering distributional aspects in evaluating aid effectiveness and by providing policy relevant insights for improving the targeting of aid programs.

2. Literature Review

2.1. Introduction to Poverty Measurement

The conceptualization and measurement of poverty have undergone significant evolution over the past several decades, reflecting broader shifts in development thinking and empirical methodology. Early approaches to poverty measurement were primarily income-based and focused on identifying the proportion of individuals living below a predefined poverty line (Kaufmann et al. 2010). The most widely used measure in this regard has been the poverty headcount ratio, which provides a simple and intuitive representation of poverty incidence (Chakravarty, 1983). While this measure has been extensively used in empirical research and policy analysis, it has been increasingly criticized for its inability to capture the depth and distribution of poverty (Mosley et al. 2004).

One of the key limitations of the headcount ratio is that it treats all individuals below the poverty line as equally poor, regardless of how far their incomes fall below that threshold. As a result, it fails to distinguish between individuals who are marginally poor and those who are severely deprived. This limitation has important implications for policy evaluation, as interventions that lift individuals just above the poverty line may significantly reduce the headcount ratio without addressing deeper forms of deprivation. Consequently, reliance on this measure alone may lead to an incomplete or even misleading assessment of poverty dynamics.

In response to these limitations, researchers have developed more comprehensive measures of poverty that account for both the depth and severity of deprivation. Among these, the Foster-Greer-Thorbecke (FGT) class of poverty measures has become the standard framework in empirical poverty analysis. The FGT measures provide a flexible and decomposable approach that allows researchers to capture multiple dimensions of poverty using a single parameterized formula. By varying the parameter, it is possible to obtain measures of poverty incidence, depth, and severity, thereby offering a more nuanced understanding of poverty dynamics (Foster et al. 1984; Ravallion, 2016).

The adoption of multidimensional poverty measures has significantly improved the ability of researchers to analyze the distributional effects of economic policies and external interventions. In particular, these measures are well-suited for examining the impact of foreign aid, as they allow for a more detailed assessment of how aid affects different segments of the population. This study builds on this framework by employing the FGT indices to evaluate whether foreign aid effectively reaches the poorest individuals in Pakistan.

2.2. Aid and Poverty Reduction

The relationship between foreign aid and poverty reduction has been a central focus of development economics, with a large body of empirical literature examining the effectiveness of aid in improving welfare outcomes (World Bank, 2023). Early studies on aid effectiveness were primarily concerned with its impact on economic growth, based on the assumption that growth would lead to poverty reduction through a trickle-down effect (Sen, 1981; Sen 1987). However, the mixed and often inconclusive findings in the aid-

growth literature have led researchers to question the validity of this indirect transmission mechanism.

As a result, more recent studies have shifted attention toward the direct impact of foreign aid on poverty. This approach recognizes that aid can influence poverty through targeted interventions such as social protection programs, health services, and education initiatives. Alvi and Senbeta (2011) provide one of the most comprehensive analyses of the direct relationship between aid and poverty, finding that foreign aid has a statistically significant negative effect on poverty. However, they also note that the magnitude of this effect varies across countries and depends on the inclusion of control variables such as income and inequality.

Other studies have produced similar findings, highlighting the potential of foreign aid to reduce poverty under certain conditions. For example, Mosley and Suleiman (2007) argue that aid is more effective when it is allocated to sectors that directly benefit the poor, such as agriculture and rural development. Similarly, Gomanee et al. (2005) find that aid can improve social indicators, including health and education outcomes, which are closely linked to poverty reduction.

Despite these positive findings, the literature also highlights several challenges associated with aid effectiveness. One of the most prominent concerns is that aid may be poorly targeted and may not reach the intended beneficiaries (Beck et al. 2000). Boone (1996) argues that a significant portion of aid is captured by political elites and does not benefit the poor. Similarly, Easterly et al. (2004) emphasize the role of policy environments and governance structures in shaping the effectiveness of aid.

These conflicting findings suggest that the relationship between foreign aid and poverty is complex and context dependent. While aid has the potential to reduce poverty, its effectiveness depends on a range of factors, including how it is allocated, implemented, and monitored. This study contributes to this literature by examining not only the overall impact of aid on poverty but also its distributional effects across different dimensions of poverty.

2.3. Distributional Effects of Aid

A growing body of literature has emphasized the importance of distributional effects in evaluating the impact of foreign aid. While aggregate measures of poverty provide useful information about overall trends, they may mask important differences in how benefits are distributed across the population. In particular, there is increasing concern that foreign aid may disproportionately benefit individuals who are closer to the poverty line, while having limited impact on those who are most severely deprived.

Ravallion (2001) argues that economic growth does not automatically translate into poverty reduction, particularly in the presence of high levels of inequality. This insight has important implications for the evaluation of foreign aid, as it suggests that aggregate improvements in poverty may not reflect improvements for the poorest segments of society. Similarly, Dollar and Kraay (2002) find that while growth tends to benefit the poor on

average, the distribution of these benefits can vary significantly across countries and time periods.

In the context of foreign aid, these findings suggest that the effectiveness of aid should be evaluated not only in terms of its impact on overall poverty levels but also in terms of its distributional effects. Mosley and Suleiman (2007) provide evidence that aid is more effective in reducing moderate poverty than extreme poverty, indicating that the poorest individuals may not fully benefit from aid programs. This may be due to a variety of factors, including targeting inefficiencies, institutional constraints, and structural barriers that limit access to resources.

Another important consideration is the role of inequality in shaping the distributional impact of aid. High levels of inequality can reduce the effectiveness of aid by concentrating resources among a small segment of the population. In such contexts, even large inflows of aid may have limited impact on poverty reduction, as the benefits are not evenly distributed.

This study builds on this literature by explicitly examining the distributional effects of foreign aid using multidimensional poverty measures. By comparing the impact of aid across the headcount ratio, poverty gap, and poverty severity indices, the study provides new insights into how aid affects different segments of the population.

2.4. Multidimensional Poverty and the FGT Framework

The Foster-Greer-Thorbecke (FGT) framework represents one of the most significant advancements in the measurement of poverty, providing a flexible and theoretically grounded approach for capturing its multiple dimensions. The FGT indices are defined by a parameter that determines the sensitivity of the measure to the depth and severity of poverty. When the parameter is set to zero, the measure corresponds to the poverty headcount ratio, which captures the incidence of poverty. When the parameter is set to one, the measure becomes the poverty gap index, which reflects the average shortfall of income from the poverty line. When the parameter is set to two, the measure captures poverty severity, giving greater weight to individuals who are further below the poverty line (Atkinson, 1987; Bourguignon, 2003).

One of the key advantages of the FGT framework is its decomposability, which allows researchers to analyze poverty at different levels of aggregation and to examine the contribution of different groups to overall poverty. This feature is particularly useful for policy analysis, as it enables the identification of specific populations or regions that are most affected by poverty (Datt & Ravallion, 1992; Deaton, 1997).

The application of the FGT framework in the context of foreign aid provides valuable insights into the distributional impact of aid. By examining how aid affects different dimensions of poverty, researchers can assess whether aid programs are effectively targeting the poorest individuals or whether they are primarily benefiting those who are less deprived.

Despite the extensive literature on foreign aid and poverty, several important gaps remain. First, most studies focus on aggregate measures of poverty, such as the headcount ratio,

and do not consider the depth and severity of poverty. This limits the ability to assess the distributional impact of aid and may lead to incomplete conclusions.

Second, while some studies have examined the direct impact of aid on poverty, relatively few have employed multidimensional poverty measures within a dynamic panel framework. This represents a significant gap; as such an approach is necessary to capture both the persistence of poverty and the distributional effects of aid.

Third, there is limited empirical evidence on the distributional impact of foreign aid in the context of Pakistan. Given the country's unique economic and institutional characteristics, this represents an important area for further research.

This study addresses these gaps by providing a comprehensive analysis of the impact of foreign aid on multiple dimensions of poverty using a dynamic panel data approach. By doing so, it contributes to a more nuanced understanding of aid effectiveness and provides valuable insights for policymakers and researchers.

3. Methods

3.1. Conceptual Framework

The conceptual foundation of this study is rooted in the recognition that poverty is a multidimensional and distribution-sensitive phenomenon, and that the impact of foreign aid on poverty cannot be adequately captured through aggregate measures alone. Traditional approaches to evaluating aid effectiveness often rely on the poverty headcount ratio, implicitly assuming that reductions in poverty incidence reflect uniform improvements across all segments of the population (Kakwani, 1993). However, this assumption overlooks the heterogeneity of poverty and the possibility that policy interventions may disproportionately benefit certain groups while leaving others behind. This study challenges this assumption by explicitly incorporating the distributional dimension of poverty into the analysis (Kanbur, 2001).

At the core of the conceptual framework is the distinction between poverty incidence, depth, and severity, as captured by the FGT class of poverty measures. Poverty incidence (P_0) reflects the proportion of individuals living below the poverty line, while poverty depth (P_1) measures the average shortfall of income from the poverty line, and poverty severity (P_2) captures inequality among the poor by assigning greater weight to those who are further below the threshold. These measures collectively provide a comprehensive representation of poverty and allow for a more nuanced assessment of how foreign aid affects different segments of the population (Foster et al. 1984).

Foreign aid is conceptualized as an external resource flow that can influence poverty through multiple channels. These include direct transfers to households, public expenditure on social services, infrastructure development, and investments in human capital. In theory, such interventions should reduce poverty by increasing incomes, improving access to essential services, and enhancing economic opportunities. However, the effectiveness of these channels depends critically on how resources are allocated and distributed within the economy (Rajan & Subramanian, 2011; Kosack, 2003; Rodrik et al. 2004).

A key insight of this study is that the impact of foreign aid is not uniform across the poverty distribution. Individuals who are close to the poverty line may benefit more readily from aid-funded interventions, as relatively small increases in income or access to services can lift them above the threshold. In contrast, individuals who are deeply impoverished may require more substantial and sustained interventions to achieve similar outcomes. As a result, aid may have a stronger impact on poverty incidence than on poverty depth or severity.

This asymmetry in impact can be understood through the lens of marginal returns to aid. For individuals near the poverty line, the marginal benefit of aid is relatively high, as small changes can lead to significant improvements in their economic status. For those at the lower end of the distribution, however, the marginal benefit may be lower due to structural constraints such as lack of access to education, healthcare, and financial services. These constraints limit the ability of the poorest individuals to fully benefit from aid, resulting in a weaker impact on poverty depth and severity.

Another important aspect of the conceptual framework is the role of inequality in shaping the distributional effects of aid (Levine, 2005). High levels of inequality can reduce the effectiveness of aid by concentrating resources among a small segment of the population. In such contexts, even large inflows of aid may have limited impact on poverty reduction if the benefits are not evenly distributed. This highlights the importance of considering not only the magnitude of aid but also its distribution across different groups.

The conceptual framework also recognizes the dynamic nature of poverty. Poverty is not a static condition but evolves over time due to changes in economic conditions, policy interventions, and external shocks. The inclusion of lagged poverty in the empirical model reflects this dynamic behavior and captures the persistence of poverty. This persistence is particularly relevant in developing countries, where structural factors such as limited access to resources and opportunities can trap individuals in long-term poverty.

Overall, the conceptual framework of this study emphasizes the importance of adopting a multidimensional and dynamic approach to analyzing the impact of foreign aid on poverty. By explicitly accounting for differences in poverty incidence, depth, and severity, the study provides a more comprehensive understanding of how aid affects different segments of the population and highlights the need for distribution-sensitive policy interventions.

3.2. Empirical Framework

Building on the conceptual framework, the empirical strategy of this study is designed to estimate the impact of foreign aid on multiple dimensions of poverty while addressing key econometric challenges such as endogeneity, persistence, and unobserved heterogeneity. The analysis is conducted using a dynamic panel data model, which allows for the incorporation of lagged dependent variables and captures the temporal evolution of poverty.

The baseline empirical model is specified separately for each poverty measure $k \in \{0, 1, 2\}$, corresponding to poverty incidence, depth, and severity:

$$\log P_{it}^{(k)} = \alpha_0 + \alpha_1 \log P_{i,t-1}^{(k)} + \alpha_2 \log AID_{it} + \alpha_3 \log Y_{it} + \alpha_4 \log G_{it} + \alpha_5 Z_{it} + \mu_i + \epsilon_{it} \quad (1)$$

Where $P_{it}^{(k)}$ represents the poverty measure for region i at time t , AID_{it} denotes foreign aid as a percentage of GDP, Y_{it} is GDP per capita, G_{it} represents income inequality, and Z_{it} is a vector of additional control variables, including trade openness and financial development. The term μ_i captures unobserved time-invariant heterogeneity, while ϵ_{it} is the idiosyncratic error term.

The inclusion of the lagged dependent variable $\log P_{i,t-1}^{(k)}$ introduces a dynamic component to the model, capturing the persistence of poverty over time. This is particularly important in the context of developing economies, where structural constraints and limited mobility can result in long-term poverty traps (Bourguignon, 2003; Baltagi, 2005). The coefficient α_1 is expected to be positive and less than one, indicating partial persistence.

The coefficient α_2 captures the elasticity of poverty with respect to foreign aid and represents the primary parameter of interest. A negative value of α_2 indicates that foreign aid reduces poverty, while differences in the magnitude of this coefficient across the three models provide insights into the distributional effects of aid. Specifically, a larger absolute value for the headcount model relative to the gap and severity models would suggest that aid is more effective in reducing poverty incidence than in addressing deeper forms of poverty.

A major econometric challenge in estimating this model is the potential endogeneity of foreign aid and income variables (Baltagi, 2005; Mosley et al. 2004). Foreign aid is often allocated based on the economic conditions of recipient countries, leading to reverse causality. Similarly, GDP per capita may be jointly determined with poverty, resulting in simultaneity bias. These issues can lead to biased and inconsistent estimates if not properly addressed.

To overcome these challenges, this study employs the System Generalized Method of Moments (System GMM) estimator developed by Arellano and Bover (1995) and Blundell and Bond (1998). This estimator is particularly well-suited for dynamic panel data models with endogenous regressors and unobserved heterogeneity. It combines equations in first differences and levels, using lagged values of endogenous variables as instruments.

The differenced equation is given by:

$$\Delta \log P_{it}^{(k)} = \alpha_1 \Delta \log P_{i,t-1}^{(k)} + \alpha_2 \Delta \log AID_{it} + \alpha_3 \Delta \log Y_{it} + \alpha_4 \Delta \log G_{it} + \Delta \epsilon_{it} \quad (2)$$

This transformation eliminates the fixed effects μ_i , while lagged levels of the variables are used as instruments for the differenced regressors. The level equation, on the other hand, uses lagged differences as instruments, improving efficiency and reducing finite sample bias.

The validity of the instruments is a critical aspect of the estimation strategy. To assess this, the study employs the Hansen test of overidentifying restrictions, which evaluates whether the instruments are uncorrelated with the error term. A failure to reject the null hypothesis

indicates that the instruments are valid. In addition, the Arellano-Bond test for autocorrelation is used to examine the serial correlation of the error term. While first-order autocorrelation is expected, the absence of second-order autocorrelation is necessary for the consistency of the estimator.

To ensure robustness, the model is estimated using alternative specifications, including the exclusion of certain control variables and the use of different lag structures. The results are also compared with those obtained from fixed effects models to assess consistency.

Overall, the empirical framework provides a rigorous approach for analyzing the impact of foreign aid on multiple dimensions of poverty. By combining a dynamic panel model with System GMM estimation, the study addresses key econometric challenges and provides reliable estimates of the distributional effects of foreign aid.

4. Results

4.1. Descriptive Statistics

The descriptive statistics provide a foundational understanding of the distributional characteristics of poverty and key explanatory variables over the study period from 1995 to 2025. Given the focus of this study on multidimensional poverty, particular attention is paid to the variation across the three FGT indices, which capture the incidence, depth, and severity of poverty.

The statistics reveal several important patterns. First, poverty remains persistent across all dimensions, with the headcount ratio averaging above 30 percent. However, the relatively lower mean values of the poverty gap and severity indices suggest that while a significant portion of the population is classified as poor, the average intensity of deprivation is comparatively moderate. This discrepancy highlights the importance of distinguishing between different dimensions of poverty, as aggregate measures may obscure underlying inequalities.

Table 1: Descriptive Statistics (Panel Data: 1995-2025)

Variable	Mean	Std. Dev	Min	Max
Poverty Headcount (P_0)	31.48	8.05	18.20	61.10
Poverty Gap (P_1)	7.02	2.18	4.90	15.60
Poverty Severity (P_2)	2.19	0.88	1.10	5.50
Foreign Aid (% GDP)	2.29	0.79	0.70	4.20
GDP per capita (log)	7.84	0.42	6.90	8.70
Gini Coefficient	0.32	0.05	0.24	0.41
Financial Development	22.75	4.68	13.80	31.60
Trade Openness	0.34	0.09	0.18	0.55

Second, the variability in foreign aid flows indicates fluctuations in external support over time, reflecting changes in global economic conditions and donor priorities. The relatively narrow range of aid suggests that while aid is a consistent feature of the economy, its magnitude is not sufficiently large to generate transformative effects without complementary policies.

Third, the variation in income inequality and financial development underscores the presence of structural heterogeneity across regions and time. These factors are likely to influence how the benefits of aid are distributed, thereby affecting its impact on different poverty measures.

4.2. System GMM Regression Results

The core empirical findings are derived from the System GMM estimation, which accounts for endogeneity, persistence, and unobserved heterogeneity. The results provide a comprehensive assessment of how foreign aid affects different dimensions of poverty.

Table 2: *System GMM Results (Multidimensional Poverty Models)*

Variable	P ₀ (Headcount)	P ₁ (Gap)	P ₂ (Severity)
Lagged Poverty	0.73***	0.69***	0.65***
Foreign Aid	-4.52***	-2.68**	-1.91**
GDP per capita	-3.28***	-3.84***	-3.46**
Gini Coefficient	2.56*	2.18*	3.12**
Financial Development	-1.15**	-1.02**	-0.85*
Trade Openness	-0.88	-0.64	-0.51

(*, **, *** denote significance at 10%, 5%, and 1%)

The coefficient on lagged poverty is positive and highly significant across all specifications, confirming the strong persistence of poverty. This result reflects the structural nature of poverty in developing economies, where limited access to resources and opportunities creates long-term constraints on upward mobility.

The coefficient on foreign aid is negative and statistically significant across all poverty measures, indicating that aid contributes to poverty reduction. However, the magnitude of the coefficient varies substantially across the three models. The effect is strongest for the headcount index, suggesting that aid is particularly effective in reducing the number of individuals living below the poverty line. In contrast, the smaller coefficients for the poverty gap and severity indices indicate that aid is less effective in addressing deeper forms of poverty.

This pattern provides critical evidence for the central argument of this study: foreign aid does not equally benefit all segments of the poor. Specifically, aid appears to be more effective in lifting individuals who are close to the poverty line, while having a limited impact on those who are further below it. This asymmetry can be attributed to several factors.

First, aid programs may be designed in a way that prioritizes easily measurable outcomes, such as reducing the headcount ratio, rather than addressing structural inequalities. Second, individuals who are moderately poor may be better positioned to take advantage of aid-funded opportunities, such as employment or access to services, compared to those who face more severe constraints. Third, targeting inefficiencies and institutional limitations may prevent aid from reaching the most vulnerable populations.

The control variables behave as expected. GDP per capita has a strong negative effect on poverty, confirming the role of economic growth in reducing poverty. However, the persistence of significant aid effects even after controlling for income suggests that aid operates through direct channels. Income inequality has a positive and significant effect,

particularly for poverty severity, indicating that inequality exacerbates deeper forms of poverty.

4.3. Instrument Validity and Diagnostic Tests

To ensure the reliability of the System GMM estimates, several diagnostic tests are conducted.

Table 3: *Instrument Validity and Diagnostic Tests*

Test	P ₀ Model	P ₁ Model	P ₂ Model
Hansen Test (p-value)	0.35	0.42	0.39
Sargan Test (p-value)	0.29	0.37	0.34
AR(1) (p-value)	0.02	0.03	0.04
AR(2) (p-value)	0.48	0.53	0.50
Number of Instruments	27	27	27

The Hansen and Sargan tests indicate that the instruments used in the estimation are valid, as the null hypothesis of instrument exogeneity cannot be rejected. The Arellano-Bond test confirms the absence of second-order autocorrelation, ensuring the consistency of the estimator. These results provide strong support for the robustness of the empirical findings.

4.4. Robustness and Sensitivity Analysis

To further validate the results, several robustness checks are performed.

Table 4: *Robustness and Sensitivity Analysis*

Specification	Aid Effect (P0)	Aid Effect (P1)	Aid Effect (P2)
Baseline Model	-4.52***	-2.68**	-1.91**
Without GDP Control	-4.20***	-2.50**	-1.80**
Fixed Effects Model	-3.60**	-2.10**	-1.50*
Random Effects Model	-3.85**	-2.30**	-1.60*
Pre-COVID Sample	-4.40***	-2.60**	-1.85**
Post-COVID Sample	-4.80***	-2.90**	-2.05**

The results remain consistent across different specifications, confirming that the findings are not driven by model assumptions or sample selection. The slightly stronger effects in the post-COVID period suggest that aid may have played a stabilizing role during economic disruptions.

The results provide strong and consistent evidence that foreign aid reduces poverty in Pakistan, but its impact varies significantly across different dimensions of poverty. The stronger effect on poverty incidence compared to poverty depth and severity indicates that aid is more effective in addressing moderate poverty than extreme poverty.

This finding has important implications for the evaluation of aid effectiveness. While reductions in the headcount ratio may suggest progress, they may not reflect improvements for the poorest segments of the population. The weaker impact on poverty severity suggests that structural constraints and targeting inefficiencies limit the ability of aid to reach the most vulnerable individuals.

Overall, the results support the central argument of this study: foreign aid does not fully reach the poorest, highlighting the need for more targeted and distribution-sensitive policy interventions.

5. Discussion and Conclusion

The empirical findings of this study provide a comprehensive and nuanced understanding of the relationship between foreign aid and poverty, particularly when poverty is examined through a multidimensional and distribution-sensitive framework. By employing the FGT indices, the analysis moves beyond conventional approaches that focus solely on poverty incidence and instead captures the depth and severity of deprivation. This methodological choice reveals critical insights that would otherwise remain obscured in aggregate analyses.

At a fundamental level, the results confirm that foreign aid has a statistically significant and economically meaningful impact on poverty reduction. However, the magnitude of this effect varies considerably across different dimensions of poverty. The strongest effect is observed for the poverty headcount index, indicating that foreign aid is effective in reducing the proportion of individuals living below the poverty line. In contrast, the effects on the poverty gap and severity indices are comparatively weaker, suggesting that aid is less effective in addressing deeper and more entrenched forms of poverty. This divergence across poverty measures constitutes the central finding of the study and has important implications for both theory and policy.

From a theoretical perspective, the results challenge the implicit assumption in much of the aid effectiveness literature that poverty reduction is a homogeneous process. Instead, the findings highlight the importance of considering heterogeneity within the poor population. Individuals who are close to the poverty line may require relatively modest improvements in income or access to services to escape poverty, making them more responsive to aid-funded interventions. In contrast, individuals who are deeply impoverished face multiple and overlapping constraints, including limited access to education, healthcare, financial services, and employment opportunities. These constraints reduce the marginal effectiveness of aid and make it more difficult for such individuals to benefit from external interventions.

This asymmetry in the impact of foreign aid can also be understood in terms of the structure and design of aid programs. Many aid initiatives are oriented toward achieving measurable and short-term outcomes, such as reducing poverty incidence or improving access to basic services. While these objectives are important, they may inadvertently prioritize interventions that benefit individuals who are easier to reach, such as those living in urban or semi-urban areas or those with some existing level of economic participation. As a result, the poorest and most marginalized individuals, who often reside in remote or underserved regions, may be excluded from the benefits of aid.

The findings also highlight the role of structural inequality in shaping the distributional impact of foreign aid. The positive and significant relationship between income inequality and poverty severity suggests that unequal distribution of resources exacerbates deeper forms of deprivation. In such contexts, foreign aid may be insufficient to offset the effects of inequality, particularly if it is not explicitly targeted toward the most vulnerable groups. This underscores the importance of integrating distributional considerations into the design and implementation of aid programs.

Another important implication of the results relates to the persistence of poverty, as indicated by the strong and significant coefficient on the lagged dependent variable. This finding reflects the presence of poverty traps, where individuals remain stuck in cycles of deprivation due to structural barriers. The persistence of poverty implies that short-term interventions may be insufficient to achieve lasting improvements, and that sustained and targeted efforts are required to address the underlying causes of poverty.

The results also provide important insights into the role of economic growth and inequality in shaping poverty outcomes. While GDP per capita is found to have a strong negative effect on poverty, indicating that growth contributes to poverty reduction, this effect is not sufficient to eliminate deeper forms of deprivation. Similarly, the positive effect of inequality suggests that the benefits of growth are not evenly distributed, limiting its impact on poverty severity. These findings reinforce the argument that growth alone is not sufficient to achieve inclusive development and that targeted interventions are necessary to ensure that the benefits of growth and aid reach the poorest segments of the population.

From a policy perspective, the findings of this study have several important implications. First, they highlight the need for a shift in focus from aggregate poverty reduction to distribution-sensitive approaches that explicitly address the needs of the poorest individuals. Policymakers should prioritize interventions that target extreme poverty, such as conditional cash transfer programs, rural development initiatives, and investments in basic infrastructure in underserved areas. These interventions should be designed to address the specific constraints faced by the poorest populations, including limited access to education, healthcare, and financial services.

Second, the results underscore the importance of improving the targeting mechanisms of aid programs. Effective targeting requires accurate identification of vulnerable populations and the use of delivery mechanisms that ensure resources reach intended beneficiaries. This may involve the use of data-driven approaches, community-based targeting, and digital technologies to improve transparency and accountability.

Third, the findings suggest that foreign aid should be complemented by domestic policies aimed at reducing inequality and promoting inclusive growth. Without such complementary measures, the impact of aid on poverty reduction may be limited, particularly for the poorest segments of the population. Policies that promote equitable access to resources, such as education and financial services, can enhance the effectiveness of aid and contribute to more inclusive development outcomes.

Fourth, the study highlights the importance of long-term and sustained interventions in addressing poverty. Given the persistence of poverty, short-term aid programs may not be sufficient to achieve lasting improvements. Instead, policymakers and donors should focus on designing programs that provide continuous support and address the structural causes of poverty.

The findings also have implications for international donors, who play a critical role in shaping the allocation and design of aid programs. Donors should prioritize not only the volume of aid but also its effectiveness and distributional impact. This may involve adopting a more targeted approach to aid allocation, focusing on sectors and regions where

the need is greatest and where the potential for impact is highest. In addition, donors should support efforts to improve governance and institutional capacity in recipient countries, as these factors play a crucial role in determining the effectiveness of aid.

In the context of Pakistan, the results of this study are particularly relevant given the country's persistent poverty challenges and regional disparities. The findings suggest that while foreign aid has contributed to reducing overall poverty, it has not been sufficient to address deeper forms of deprivation. This highlights the need for a more targeted and inclusive approach to development policy, one that prioritizes the needs of the most vulnerable populations.

Despite its contributions, this study also acknowledges certain limitations that provide opportunities for future research. One limitation is the use of aggregate data, which may mask heterogeneity at the household level. Future studies could use micro-level data to examine how aid affects different groups within the population. Additionally, further research could explore the role of other factors, such as institutional quality and financial development, in shaping the distributional impact of aid.

In conclusion, this study provides strong empirical evidence that foreign aid reduces poverty but does not equally benefit all segments of the population. While aid is effective in reducing poverty incidence, its impact on poverty depth and severity is limited, indicating that the poorest individuals are not fully reached. This finding has important implications for the design and evaluation of aid programs, highlighting the need for more targeted and distribution-sensitive approaches. Achieving meaningful and inclusive poverty reduction requires not only financial resources but also a deep understanding of the structural factors that shape poverty and the distributional dynamics of policy interventions.

Statements and Declarations

We, the author of this research paper declare no competing interests. This paper is totally academic in nature where no funding was received from any person or institution. Moreover, data is collected from secondary sources, and will be furnished upon request.

References

- Alvi, E., and Senbeta, A. (2012). Does foreign aid reduce poverty? *Journal of International Development*, 24(8), 955–976. <<https://doi.org/10.1002/jid.1790>>
- Arellano, M., and Bover, O. (1995). Another look at the instrumental variable estimation of error-components models. *Journal of Econometrics*, 68(1), 29–51. <[https://doi.org/10.1016/0304-4076\(94\)01642-D](https://doi.org/10.1016/0304-4076(94)01642-D)>
- Atkinson, A.B. (1987). On the measurement of poverty. *Econometrica*, 55(4), 749–764. <<https://www.jstor.org/stable/1911028>>

- Baltagi, B.H. (2005). *Econometric analysis of panel data* (3rd ed.). Wiley.
- Beck, T., Demirgüç-Kunt, A., and Levine, R. (2007). Finance, inequality and the poor. *Journal of Economic Growth*, 12(1), 27–49. <<https://doi.org/10.1007/s10887-007-9010-6>>
- Beck, T., Levine, R., and Loayza, N. (2000). Finance and the sources of growth. *Journal of Financial Economics*, 58(1–2), 261–300. <[https://doi.org/10.1016/S0304-405X\(00\)00072-6](https://doi.org/10.1016/S0304-405X(00)00072-6)>
- Blundell, R., and Bond, S. (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics*, 87(1), 115–143. <[https://doi.org/10.1016/S0304-4076\(98\)00009-8](https://doi.org/10.1016/S0304-4076(98)00009-8)>
- Boone, P. (1996). Politics and the effectiveness of foreign aid. *European Economic Review*, 40(2), 289–329. <[https://doi.org/10.1016/0014-2921\(95\)00127-1](https://doi.org/10.1016/0014-2921(95)00127-1)>
- Bourguignon, F. (2003). The growth elasticity of poverty reduction. World Bank. <<http://documents.worldbank.org/curated/en/503161468780002293>>
- Burnside, C., and Dollar, D. (2000). Aid, policies, and growth. *American Economic Review*, 90(4), 847–868. <<https://doi.org/10.1257/aer.90.4.847>>
- Chakravarty, S.R. (1983). A new index of poverty. *Mathematical Social Sciences*, 6(3), 307–313. <[https://doi.org/10.1016/0165-4896\(83\)90064-1](https://doi.org/10.1016/0165-4896(83)90064-1)>
- Collier, P., and Dollar, D. (2002). Aid allocation and poverty reduction. *European Economic Review*, 46(8), 1475–1500. <[https://doi.org/10.1016/S0014-2921\(01\)00187-8](https://doi.org/10.1016/S0014-2921(01)00187-8)>
- Datt, G., and Ravallion, M. (1992). Growth and redistribution components of changes in poverty measures. *Journal of Development Economics*, 38(2), 275–295. <[https://doi.org/10.1016/0304-3878\(92\)90001-P](https://doi.org/10.1016/0304-3878(92)90001-P)>
- Deaton, A. (1997). The analysis of household surveys. World Bank. <<http://documents.worldbank.org/curated/en/593871468777303124>>
- Dollar, D., and Kraay, A. (2002). Growth is good for the poor. *Journal of Economic Growth*, 7(3), 195–225. <<https://doi.org/10.1023/A:1020139631000>>
- Dollar, D., and Kraay, A. (2004). Trade, growth, and poverty. *Economic Journal*, 114(493), F22–F49. <<https://doi.org/10.1111/j.0013-0133.2004.00186.x>>
- Easterly, W. (2006). *The white man's burden*. Penguin Press.
- Easterly, W., Levine, R., and Roodman, D. (2004). Aid, policies, and growth. *American Economic Review*, 94(2), 774–780. <<https://doi.org/10.1257/0002828041464560>>
- Foster, J., Greer, J., and Thorbecke, E. (1984). A class of decomposable poverty measures. *Econometrica*, 52(3), 761–766. <<https://doi.org/10.2307/1913475>>
- Gomanee, K., Girma, S., and Morrissey, O. (2005). Aid, public spending and human welfare. *Journal of International Development*, 17(8), 1055–1076. <<https://doi.org/10.1002/jid.1163>>
- Haughton, J., and Khandker, S. (2009). *Handbook on poverty and inequality*. World Bank. <<http://documents.worldbank.org/curated/en/488081468157174849>>
- Kanbur, R. (2001). Economic policy, distribution and poverty. *World Development*, 29(6), 1083–1094. <[https://doi.org/10.1016/S0305-750X\(01\)00017-1](https://doi.org/10.1016/S0305-750X(01)00017-1)>
- Kakwani, N. (1993). Poverty and economic growth with application to Côte d'Ivoire. *Review of Income and Wealth*, 39(2), 121–139. <<https://doi.org/10.1111/j.1475-4991.1993.tb00443.x>>
- Kaufmann, D., Kraay, A., and Mastruzzi, M. (2010). Governance matters. World Bank Policy Research Working Paper. <<http://documents.worldbank.org/curated/en/665731468739470954>>
- Kosack, S. (2003). Effective aid. *World Development*, 31(1), 1–22. <[https://doi.org/10.1016/S0305-750X\(02\)00177-8](https://doi.org/10.1016/S0305-750X(02)00177-8)>

- Levine, R. (2005). Finance and growth. *Handbook of Economic Growth*. <[https://doi.org/10.1016/S1574-0684\(05\)01012-9](https://doi.org/10.1016/S1574-0684(05)01012-9)>
- Mosley, P., Hudson, J., and Verschoor, A. (2004). Aid, poverty reduction and the new conditionality. *Economic Journal*, 114(496), F217–F243. <<https://doi.org/10.1111/j.1468-0297.2004.00220.x>>
- Mosley, P., and Suleiman, A. (2007). Aid, Agriculture and Poverty in Developing Countries. *Review of Development Economics*, 11, 139–158. <<https://doi.org/10.1111/j.1467-9361.2006.00354.x>>
- Rajan, R.G., and Subramanian, A. (2011). Aid and Dutch disease. *Journal of Development Economics*, 94(1), 106–118. <<https://doi.org/10.1016/j.jdeveco.2009.12.004>>
- Ravallion, M. (2001). Growth, inequality and poverty. World Bank. <<https://doi.org/10.1596/1813-9450-2558>>
- Ravallion, M. (2016). *The economics of poverty*. Oxford University Press.
- Rodrik, D., Subramanian, A. and Trebbi, F. (2004). Institutions Rule: The Primacy of Institutions Over Geography and Integration in Economic Development. *Journal of Economic Growth*, 9, 131–165. <<https://doi.org/10.1023/B:JOEG.0000031425.72248.85>>
- Sen, A. (1981). *Poverty and famines*. Oxford University Press.
- Sen, A. (1987). *The standard of living*. Cambridge University Press.
- World Bank. (2023). World development indicators. <<https://doi.org/10.1596/978-1-4648-0163-1>>

Mathematical modeling of nearshoring and reshoring in the reconfiguration of trade flows

Rares NITU

Bucharest Academy of Economic Studies, Romania
raresmnitu@gmail.com

Abstract. *The paper develops an integrated mathematical framework for identifying and evaluating nearshoring and reshoring, two processes frequently associated in economic discourse, but analytically distinct. Nearshoring is treated as a redistribution of trade mass to partners located at shorter economic distances, and reshoring as an increase in the share of domestic supply in domestic absorption, conditional on the effective substitution of imports. Starting from a bilateral-sectoral-temporal tensor of flows, the analysis builds probability measures, proximity, concentration and entropy indicators, exact decompositions of the trade distance change, as well as a convex optimization problem of the supplier structure. Results are formulated and demonstrated regarding the monotony of trade transfer, the separation of geographical and sectoral effects, the necessary and sufficient conditions of reshoring and the uniqueness of the optimal solution. The framework is complemented by a PPML gravity specification, a dynamic adjustment model and a procedure for identifying structural breaks. The result is a quantitative architecture in which geographical proximity, production recovery, diversification and resilience are measured separately, but can be brought together in a formal decision rule.*

Keywords: nearshoring, reshoring, commercial distance, optimal transport, entropy, convex optimization, PPML, structural ruptures.

JEL Classification: F14, F15, F17, F23.

1. Introduction

The reconfiguration of international production chains has brought the decision of localization back to the center of economic analysis. In its simplified form, this decision is presented as a choice between production in an area with low unit costs and production in an area close to the final market. Such representation is insufficient, because the wage cost does not exhaust the total cost of supply, and the geographical distance does not fully capture time, uncertainty, infrastructure quality, geopolitical risk and operational flexibility. Moreover, the reduction of imports from a distant partner does not, in itself, demonstrate the existence of nearshoring: the flow can be replaced by an equally distant partner, it can disappear as a result of the contraction in demand or it can be absorbed by a sector with a different spatial structure. Similarly, the growth of domestic production is not automatically reshoring if it is export-oriented or if imports grow at a higher rate. The central problem is therefore one of mathematical identification: verifiable conditions must be established under which a change in the distribution of flows can be classified unambiguously (Anderson & Wincoop, 2003).

The objective of the paper is to build a model in which nearshoring and reshoring are defined by different mathematical objects. For nearshoring, bilateral and sectoral distribution of trade is necessary, as the phenomenon concerns the shift of the center of gravity of supply. For reshoring, the identity of domestic absorption is required, as the phenomenon concerns the ratio between net domestic supply and imports. The model does not aim to reduce the decision to a single index from the start; on the contrary, it separates volume, composition, proximity, concentration and risk, then shows under what conditions these dimensions can be aggregated (Tate, et al., 2014). This order is essential: a composite index constructed prior to identification may hide the fact that geographical proximity is achieved by excessive concentration or that the apparent internal return results only from decreased absorption.

The contribution is organised into five components.

The first normalizes trade flows in the form of a measure of probability and defines nearshoring by changing geographical moments.

The second demonstrates the effect of an elementary transfer of trade mass and provides an exact breakdown between the geographical reallocation within sectors and the change in sectoral shares.

The third defines reshoring by the logic of the domestic quota and derives a necessary and sufficient condition based on the difference between the growth rates of net domestic supply and imports.

The fourth formulates the choice of suppliers as a convex problem with cost, distance, risk and diversification.

The fifth transforms the theoretical construction into an estimable system through PPML, event study, and structural ruptures. Thus, mathematical analysis is not a succession of decorative formulas, but the mechanism by which concepts are defined, linked and tested.

2. Literature

The modern literature on the relocation of production starts from the revision of the classic decision of offshoring and from the extension of the notion of relevant cost. The choice of location cannot be based solely on the salary differential, as transport, safety stocks, variability of delivery times, coordination of activities, quality control and loss of flexibility can reverse the result of a simple accounting comparison. In this perspective, reshoring represents the return of a productive activity that had previously been transferred externally, and the relevant unit of analysis is not always the firm as a whole, but the product, operation, technological stage or economic function transferred. This delimitation separates the return of production from the mere expansion of domestic capacity. Consequently, a rigorous model must trace the origin of the flow, its destination, the nature of the activity and the economic function relocated, not just the aggregate variation in production (Bai & 3 Perron, 2003; Silva & Tenreiro, 2006).

Empirical studies have highlighted a set of factors that make the decision to relocate fundamentally multidimensional. Competitive pressures can favour the approximation of production to the consumer, design centres and innovation functions, especially when demand is volatile and product life cycles are short. In industries characterized by small series, customization and frequent changes in orders, the return of some activities can be explained by the need to reduce reaction time, not just by the convergence of salary costs (Martínez-Mora & Merino, 2014). However, backshoring remains selective, being more likely for products with high demands for quality, flexibility, automation and technology integration. Also, offshoring and backshoring are not perfectly symmetrical states, as the turnaround occurs after the accumulation of contractual costs, organizational routines, technological dependencies and supplier relationships that did not exist at the time of the initial move.

A second direction of the literature analyzes the organizational and institutional factors of relocation (McIvor & Bals, 2021). Determinants are typically grouped into costs, quality, time, risk, access to skills, proximity to customers, protection of knowledge and stability of the institutional framework. Reshoring can function as a mechanism to reduce risk and reconsider international engagement, without the firm abandoning internationalization; Rather, functions are redistributed between locations, depending on the strategic importance and vulnerability of each activity. A complete explanation must combine transaction costs, internal resources and capabilities, localization advantages, and production process characteristics. It follows that the optimal location depends simultaneously on the country, company, product and technological stage. From this perspective, reshoring should be modeled as a process with antecedents, contingent conditions, implementation, and outcomes, not as an isolated binary variable (Wiesmann, et al., 2017; Kinkel, 2014).

After recent global shocks, resilience has become a distinct motivation, but the literature warns against its automatic equivalence with localization. International diversification can

reduce exposure to a country-specific shock, while regional concentration can reduce transport time but amplify common risk. Therefore, proximity and resilience should be introduced separately in the objective function. This distinction is compatible with portfolio theory: the total risk depends not only on the risk of each provider, but also on the covariances between locations. A geographically close portfolio may have low logistical cost and high covariance of outages, and a dispersed one may have high cost and low covariance. It follows that nearshoring is not a universally dominant solution, but a movement on a border between cost, speed and risk (Di Mauro, et al., 2018).

Methodologically, the analysis of trade flows is based on the structural gravitational model, in which bilateral flows depend on the relative trade costs and the position of each economy vis-à-vis all partners. This structure justifies the use of large-scale fixed effects, capable of absorbing the unnoticed characteristics specific to the exporter, importer, sector and period. Flow logarithm and linear estimation can generate inconsistent coefficients in the presence of heteroscedasticity and eliminate zero-value observations; PPML estimation retains the multiplicative form, allows for the inclusion of null flows, and is suitable for dense commercial panels. Regime changes can be identified by multiple structural rupture procedures based on the overall minimization of the sum of square residues. In parallel, optimal transport theory provides a natural language for measuring the minimum cost needed to transform one geographic distribution into another, and entropic regularization makes the calculation applicable to large commercial networks (Correia, et al., 2020).

The main gap results from the incomplete separation of the phenomena. The managerial literature explains the reasons for relocation, and the econometric literature estimates flows, but less often is a system formulated in which nearshoring, reshoring, concentration and resilience are defined simultaneously and can produce divergent conclusions. The proposed framework addresses this gap with four requirements: the invariant of indicators to uniform scaling of trade; separation of geographical change from sectoral change; identification of reshoring by absorption, not by gross production; and the introduction of risk through a covariance matrix, not through an arbitrary penalty. These requirements transform analysis from a descriptive classification into a formal measurement and optimization problem (Ciabuschi, et al., 2019).

3. Mathematical analysis

3.1. Flow space and probabilistic normalisation

It is considered a home economy, partners, sectors and periods. The bilateral-sectoral trade flow is f_{ij} , and the total trade of the economy is T . For i , the elementary weight satisfies $w_i = f_{ij}/T$, so that the set of flows can be treated as a probability distribution on the Cartesian product $\mathcal{S} \times \mathcal{P}$. This normalization eliminates the effect of scale: if all flows are multiplied by the same constant c , then w_i remains unchanged. It follows that the indicators built from weights measure the structure, not

the level of trade. $i \in \mathcal{I}, j \in \mathcal{J}, k \in \mathcal{K}, t \in \mathcal{T}$ $X_{ijkt} \geq 0, X_{it} = \sum_{j,k} X_{ijkt}, X_{it} > 0, s_{ijkt} = X_{ijkt} / X_{it}, s_{ijkt} \in [0,1], \sum_{j,k} s_{ijkt} = 1, \mathcal{J} \times \mathcal{K} a > 0, s'_{ijkt} = a X_{ijkt} / (a X_{it}) = s_{ijkt}$

The share of the sector is s_{ijkt} , and the conditional share of the partner in the sector is p_{ijkt} . The relationship allows the separation of the sectoral composition from the geographical distribution within each sector. Be the economic distance between i and j , which can combine mileage, transportation time, logistics cost, and institutional risk. In order to limit the influence of extreme values, cost is used, and the average commercial distance becomes $q_{ikt} = \sum_j s_{ijkt} p_{ijkt} = s_{ijkt} / q_{ikt}, q_{ikt} > 0, s_{ijkt} = q_{ikt} p_{ijkt} d_{ij} > 0, j c_{ij} = \ln d_{ij}$

$$D_{it} = \sum_{j,k} s_{ijkt} c_{ij} = \sum_k q_{ikt} \mu_{ikt}, \quad \mu_{ikt} = \sum_j p_{ijkt} c_{ij}.$$

The first equality defines the aggregate geographic momentum, and the second shows that this is an average of sectoral distances. It follows from this identity that a decrease in it may come either from the proximity of the partners within the sectors or from the increase in the share of sectors that were already oriented towards proximity. Only the first mechanism represents actual geographical reallocation. $\mu_{ikt} D_{it}$

3.2. Definition of nearshoring and demonstration of monotony

Weak nearshoring between periods t and $t-1$ by δ and strictly nearshoring by δ , provided that the set of products analysed and the distance function are comparable. To demonstrate the meaning of this definition, it is considered an elementary mass transfer from the distant partner to the close partner, where $\delta > 0$. The weights also become $s'_{iakt} = s_{iakt} + \delta s'_{ibkt} = s_{ibkt} - \delta, 0 < \delta \leq s_{ibkt}$

$$D'_{it} - D_{it} = \delta c_{ia} - \delta c_{ib} = \delta (c_{ia} - c_{ib}) < 0.$$

H 1. Any positive transfer of trade mass from a partner with a higher distance cost to a lower cost partner strictly reduces the average trade distance, if the total volume and sector structure are kept constant. The proof is contained in the preceding relation: the sign of variation is the product of δ and the negative difference $c_{ia} - c_{ib}$. The result is linear for an elementary transfer and extends to a transfer matrix T , where it represents the mass moved from i to j . The overall variation is $\Delta D_i = \sum_{a,b} \tau_{ab} (c_{ia} - c_{ib}) c_{ia} < c_{ib}$; Therefore, a sufficient condition for nearshoring is that all active transfers satisfy $\tau_{ab} (c_{ia} - c_{ib}) \geq 0$. The condition is not necessary, as transfers to more distant partners can be offset by larger transfers to close partners. $\delta > 0, c_{ia} - c_{ib} T = (\tau_{ab}) \tau_{ab} \geq 0, \Delta D_i = \sum_{a,b} \tau_{ab} (c_{ia} - c_{ib}) c_{ia} < c_{ib}$

For comparisons between countries, the normalized proximity index is defined. If $c_i = \min_j c_{ij}$ and $c_i^{\max} = \max_j c_{ij}$, then $c_i^{\min} = \min_j c_{ij}, c_i^{\max} = \max_j c_{ij}$

$$N_{it} = \frac{c_i^{\max} - D_{it}}{c_i^{\max} - c_i^{\min}}, \quad 0 \leq N_{it} \leq 1.$$

The index is equal to one when the entire table is placed at the minimum distance and zero when it is placed at the maximum distance. Its derivative of a transference from is , which confirms the monotony. The advantage of normalization is comparability, and its limit is dependence on support: adding a very distant partner modifies even if that partner is not used. For this reason, robust analysis can use fixed percentiles of distance or a common support for all periods. $\partial N_{it} / \partial \delta = (c_{ib} - c_{ia}) / (c_i^{\max} - c_i^{\min}) > 0$

The average moment does not fully capture the cast. Two structures can have the same , but one can concentrate the flow at intermediate distances and the other at extremes. Enter the geographic variance and tail risk, where the weighted quantile is. Robust nearshoring requires not only , but also the avoidance of excessive growth . An economy can reduce the average by moving a large mass to the neighborhood, but it can maintain a critical dependence on a very distant supplier; The queue indicator detects this situation. $D_{it} V_{it} = \sum_{j,k} s_{ijkt} (c_{ij} - D_{it})^2 C_{it}(\alpha) = \sum_{j,k} s_{ijkt} c_{ij} \mathbf{1}(c_{ij} \geq Q_{it}(\alpha)) / (1 - \alpha) Q_{it}(\alpha) \Delta D_{it} < 0$

3.3. Exact breakdown of geographical change

For two periods, and , denote , and . Using the symmetric averages and , the exact identity is obtained $01 \Delta D_i = D_{i1} - D_{i0} \Delta q_{ik} = q_{ik1} - q_{ik0} \Delta \mu_{ik} = \mu_{ik1} - \mu_{ik0} \bar{q}_{ik} = (q_{ik1} + q_{ik0}) / 2 \bar{\mu}_{ik} = (\mu_{ik1} + \mu_{ik0}) / 2$

$$\Delta D_i = \sum_k \bar{q}_{ik} \Delta \mu_{ik} + \sum_k \bar{\mu}_{ik} \Delta q_{ik}.$$

The first term is the internal geographical effect of the sectors, and the second is the composition effect. The proof results from the algebraic identity applied to each sector. The development of the right side produces ; the cross-terms are canceled, remaining . By summing after the aggregate relation is obtained. $q_1 \mu_1 - q_0 \mu_0 = \bar{q}(\mu_1 - \mu_0) + \bar{\mu}(q_1 - q_0) \frac{1}{2}(q_1 \mu_1 - q_1 \mu_0 + q_0 \mu_1 - q_0 \mu_0) + \frac{1}{2}(q_1 \mu_1 - q_0 \mu_1 + q_1 \mu_0 - q_0 \mu_0) q_1 \mu_1 - q_0 \mu_0 k$

H 2. There is pure geographical nearshoring if, even if due to a positive sectoral effect. This sentence is important because it prevents the false rejection of nearshoring when the economy increases its share in a structural sector dependent on long distances. Conversely, it does not demonstrate geographical nearshoring if the first term is null or positive, and the decrease comes exclusively from the transfer of shares to sectors that are already close. In the application, the two components must be reported separately and, for comparability, can be divided by , obtaining relative contributions. $\sum_k \bar{q}_{ik} \Delta \mu_{ik} < 0 \Delta D_i \geq 0 \Delta D_i < 0 |D_{i0}|$

An additional breakdown separates the extensive margin from the intensive one. The indicator of the active link and the number of links are defined. The change in the total flow

can be written as the sum of the flows, appeared, disappeared, and persisted. For the set of persistent links, inputs, and outputs, the variation is . The same logic applies to distance-weighted mass. Thus, nearshoring can be produced by the intensification of existing close connections, the appearance of close connections or the disappearance of distant ones; The mechanisms have different implications for the productive capacity and stability of the network.

$$A_{ijkt} = \mathbf{1}(X_{ijkt} > 0)L_{it} = \sum_{j,k} A_{ijkt} \mathcal{P} \mathcal{E} \mathcal{X} \Delta X_{it} = \sum_{(j,k) \in \mathcal{P}} (X_{ijk1} - X_{ijk0}) + \sum_{(j,k) \in \mathcal{E}} X_{ijk1} - \sum_{(j,k) \in \mathcal{X}} X_{ijk0}$$

To measure the cost of complete distribution conversion, optimal transport is introduced. Let alone the geographical distributions over two periods, and the cost of moving a unit of mass from the location to . The set of feasible couplings is , and the discrete Wasserstein distance is $p^0 = (p_j^0)p^1 = (p_j^1)g_{ab}ba\Pi(p^0, p^1) = \{\Gamma \geq 0: \Gamma \mathbf{1} = p^1, \Gamma^\top \mathbf{1} = p^0\}$

$$W_g(p^0, p^1) = \min_{\Gamma \in \Pi(p^0, p^1)} \sum_{a,b} g_{ab} \Gamma_{ab}.$$

The value measures the intensity of the reconfiguration, but not the direction. The direction is given by . Therefore, a strong nearshoring type reconfiguration has simultaneously high and , while stability has small regardless of the sign of a minor variation. Entropic regularization adds term, ensuring a smooth solution and fast calculation by Sinkhorn scalings. $W_g D_{i1} - D_{i0} W_g \Delta D_i < 0 W_g \varepsilon \sum_{a,b} \Gamma_{ab} (\ln \Gamma_{ab} - 1)$

3.4. Formal definition of reshoring

Reshoring cannot be identified from the bilateral matrix alone, as it requires information about domestic production. For the sector, be it production, exports and imports. The net domestic supply destined for the national market is , and the internal absorption is . The internal supply share is defined by , with when the components are non-negative. Sector reshoring is identified by , but the logit transformation offers a stronger interpretation: $kQ_{ikt}E_{ikt}M_{ikt}S_{ikt}^D = Q_{ikt} - E_{ikt}A_{ikt} = S_{ikt}^D + M_{ikt}R_{ikt} = \frac{S_{ikt}^D}{A_{ikt0}} \leq R_{ikt} \leq 1 \Delta R_{ikt} > 0$

$$\text{logit}(R_{ikt}) = \ln \frac{R_{ikt}}{1 - R_{ikt}} = \ln S_{ikt}^D - \ln M_{ikt}.$$

Temporal differentiation results . $\Delta \text{logit}(R_{ikt}) = \Delta \ln S_{ikt}^D - \Delta \ln M_{ikt}$ **H 3.** The domestic share increases if and only if the net domestic supply grows proportionally faster than imports. The proof uses the strict monotony of the logit function on the interval : is equivalent to , and the last inequality is equivalent to . The condition also covers the situation in which both components decrease: there may be relative reshoring if imports contract faster than domestic supply, but this situation must be differentiated from

expansionary reshoring, further defined by $(0,1)\Delta R_{ikt} > 0 \Delta \logit(R_{ikt}) > 0 \Delta \ln S_{ikt}^D > \Delta \ln M_{ikt} \Delta S_{ikt}^D > 0$

Quota derivatives clarify the mechanism. For , we also get . The elasticities are also . The effect of a percentage increase in domestic supply is greater when the initial domestic share is small, as there is a large import base that can be substituted. For sector aggregation, the total share can be written , where . The aggregate variation is decomposed exactly into intra-sectoral effect and compositional effect, analogous to geographical decomposition: . Thus, the increase in the total share may come from reshoring in sectors or from an increase in the share of sectors that already had high domestic shares. $R = S/(S + M) \partial R / \partial S = M/(S + M)^2 > 0 \partial R / \partial M = -S/(S + M)^2 < 0 \partial \ln R / \partial \ln S = 1 - R \partial \ln R / \partial \ln M = -(1 - R)R_{it} = \sum_k S_{ikt}^D / \sum_k A_{ikt} R_{it} = \sum_k \omega_{ikt} R_{ikt} \omega_{ikt} = A_{ikt} / \sum_h A_{iht} \Delta R_i = \sum_k \bar{\omega}_{ik} \Delta R_{ik} + \sum_k \bar{R}_{ik} \Delta \omega_{ik}$

The separation between nearshoring and reshoring can be represented in a plane with the axes and . The dial indicates simultaneous external approach and internal return; indicates regionalisation without internal substitution; indicates an internal recovery without the approximation of residual imports; and indicates the expansion of external dependence. This representation prevents the terms from being used synonymously and allows for hybrid strategies, e.g. increasing domestic production for critical components and nearshoring for standardized components. $\Delta N_{it} \Delta R_{it} (+, +)(+, 0/-)(0/-, +)(-, -)$

3.5. Concentration, entropy and resilience

Closeness can be increased by moving flows to a single partner, which reduces distance but increases dependency. The geographical concentration is , where ; It satisfies when all partners are active. The Shannon entropy is , with , and the normalized diversification is . The relationship between the indicators is not linear, but both describe dispersion: it strongly penalizes large weights, while entropy is more sensitive to the entire support. $H_{it} = \sum_j u_{ijt}^2 u_{ijt} = \sum_k s_{ijkt} 1/J \leq H_{it} \leq 1 E_{it} = - \sum_j u_{ijt} \ln u_{ijt} 0 \leq E_{it} \leq \ln J E_{it}^* = E_{it} / \ln J H$

To measure structural change independent of distance, the Jensen–Shannon divergence is used. For distributions and , with , this is , where . The divergence is symmetrical, finite, and contained between and . A large value shows reconfiguration, but does not indicate approach; Combined with his sign, it separates the intense shift toward proximity from the mere volatility of partners.

$$\begin{aligned} u^0 u^1 m &= (u^0 + u^1) / 2 JS(u^0, u^1) = \frac{1}{2} KL(u^0 \parallel m) + \frac{1}{2} KL(u^1 \parallel m) KL(p \parallel q) \\ &= \sum_j p_j \ln(p_j / q_j) 0 \ln 2 \Delta D \end{aligned}$$

The risk of disruption is modelled by a semi-defined positive matrix, the elements of which represent the covariances of supplier shocks. For the vector of geographical weights , the

risk is . If the shocks are independent, the matrix is diagonal and diversification reduces risk by distributing the weights. If nearby suppliers are exposed to the same energy, climate or geopolitical shocks, non-diagonal elements may be large and regional concentration may increase. This provides an important mathematical conclusion: distance minimization and risk minimization coincide only under restrictive conditions regarding its structure .
 $\Sigma u \mathcal{R}(u) = u^\top \Sigma u \mathcal{R} \Sigma$

3.6. Optimisation of the supplier structure

The decision of the firm or the economy can be formulated as the choice of the vector on the simplex . Let it be the vector of total unit costs, the vector of economic distances, the risk matrix and the initial structure. Minimizes function $u = (u_1, \dots, u_j)^\top \Delta_j = \{u \geq 0: \mathbf{1}^\top u = 1\}$ $cd \Sigma u^0$

$$\min_{u \in \Delta_j} F(u) = c^\top u + \lambda d^\top u + \frac{\rho}{2} u^\top \Sigma u + \frac{\kappa}{2} \|u - u^0\|_2^2 - \eta \sum_j \ln(u_j + \epsilon),$$

where there is distance aversion, risk aversion, the cost of adjustment, and the last term favors diversification. The parameter prevents the singularity. The linear component would lead, on its own, to the allocation of the entire mass to the supplier with the lowest adjusted cost. The quadratic term of risk penalizes concentration in correlated locations, the cost of adjustment prevents unrealistic jumps, and the logarithmic barrier maintains an internal portfolio. $\lambda \rho \kappa \epsilon > 0$ $c^\top u + \lambda d^\top u$

Theorem 1. If, and , the function is strictly convex on the inside of the simplex, and the problem has a unique solution. The existence results from the continuity of function on a compact set, using the barrier extension to exclude the boundary when . Hessian is , which is positively defined because ; therefore, it is strictly convex and the minimizer is unique. $\kappa > 0, \rho \geq 0, \Sigma \succcurlyeq 0, \eta \geq 0, F\eta > 0, \nabla^2 F(u) = \rho \Sigma + \kappa I + \eta \text{diag}((u_j + \epsilon)^{-2})$ $\kappa I \succ 0$ F

The Karush–Kuhn–Tucker conditions make the connection between the parameters and the optimal weights. The Lagrangian is , with . The condition of stationarity for each is $\mathcal{L} = F(u) + v(\mathbf{1}^\top u - 1) - \zeta^\top u \zeta \geq 0$ j

$$c_j + \lambda d_j + \rho(\Sigma u)_j + \kappa(u_j - u_j^0) - \frac{\eta}{u_j + \epsilon} + v - \zeta_j = 0,$$

together with , , and . For an active supplier, , so the adjusted marginal cost must be equalized between locations. An increase of it penalizes distant suppliers; However, the exact effect on it depends on and on the cost of the adjustment. Differentiating the first-order system inside we get , where . After projecting onto the hyperplane , it results with . The formula shows that the reaction does not depend only on the order of distances, but also on the curvature induced by risk, diversification and inertia. $u_j \geq 0, \zeta_j \geq 0, \zeta_j u_j = 0$ $\sum_j$

$$u_j = 1\zeta_j = 0\lambda u\Sigma H du/d\lambda + d + \mathbf{1} dv/d\lambda = 0H = \nabla^2 F(u)\mathbf{1}^\top u = 1du/d\lambda = -P_H dP_H = H^{-1} - H^{-1}\mathbf{1}(\mathbf{1}^\top H^{-1}\mathbf{1})^{-1}\mathbf{1}^\top H^{-1}$$

In the dynamic version, the structure evolves according to , with and . The value function satisfies the Bellman equation, where it contains costs and risks. The quadratic cost explains the gradual adjustment: even if the new static optimum suddenly moves towards proximity, the dynamic solution may involve partial nearshoring over several periods. This feature is compatible with the rigidities of suppliers' capacity, contracts and certification. $u_{t+1} = u_t + a_t \mathbf{1}^\top a_t = 0u_t + a_t \geq 0V(u_t, z_t) = \min_{a_t} \{f(u_t + a_t, z_t) + \frac{\kappa}{2} \|a_t\|^2 + \beta \mathbb{E}[V(u_t + a_t, z_{t+1}) | z_t]\} z_t$

3.7. Econometric model and temporal identification

For bilateral flows, the conditional average is specified

$$\begin{aligned} \mu_{ijkt} &= \mathbb{E}(X_{ijkt} | \mathcal{F}_{ijkt}) \\ &= \exp(\alpha_{it} + \gamma_{jt} + \delta_{kt} + \phi_{ij} + \beta_1 \ln d_{ij} + \beta_2 \text{Near}_{ij} \times \text{Post}_t \\ &\quad + \beta_3 \text{Adv}_{ikt} \times \text{Near}_{ij} \times \text{Post}_t). \end{aligned}$$

The fixed export-year, importer-year and sector-year effects absorb supply, demand and technology shocks, and the pair effect controls the invariable bilateral characteristics. The variable identifies partners below a distance threshold or in the same region, marks the reconfiguration regime, and can represent the comparative advantage. The coefficient shows a relative increase in near-term flows after the change, and tests whether it is stronger in sectors where the economy has a competitive capacity. $\text{Near}_{ij} \text{Post}_t \text{Adv}_{ikt} \beta_2 > 0 \beta_3$

Poisson's pseudo-log-likelihood is , where it indexes observations. The score is , and Hessian is negatively semi-defined. The estimator satisfies the momentary condition. For a continuous variable, , and for a binary variable the exact percentage effect is . The use of PPML does not assume equality between mean and variance for the consistency of the coefficients; It is sufficient to correctly specify the conditional mean, and the inference is made with a sandwich matrix and adequate clustering. $\ell(\theta) = \sum_n [X_n \ln \mu_n - \mu_n - \ln(X_n!)]$ $nS(\theta) = \sum_n x_n (X_n - \mu_n) H(\theta) = -\sum_n \mu_n x_n x_n^\top \sum_n x_n (X_n - \hat{\mu}_n) = 0$ $\partial \ln \mu / \partial x_r = \beta_r 100(e^{\beta_r} - 1)$

To avoid arbitrary choice of period, the model can be extended to an event study. If it is time for change for the economy, it is estimated. Pre-event coefficients test pre-trends, and subsequent coefficients describe the speed and persistence of reallocation. Empirical nearshoring is sustained when the pre-event coefficients are close to zero, and the post-event coefficients become positive and persistent for close partners, at the same time as the indicator decreases. $\text{Post} T_i i \sum_{r \neq -1} \beta_r \text{Near}_{ij} \mathbf{1}(t - T_i = r) D_{it}$

Structural breaks are estimated for the indicator series or for the distance elasticity. For breaks to the data, the model is on the segment. The estimator minimizes , below a

minimum segment length. Dynamic programming identifies the global minimum, and the number of breaks can be chosen through sequential tests and informational criteria. A negative break in its level or an increase in the coefficient for close partners provides evidence of regime change, but the interpretation requires simultaneous verification of the sectoral composition and the internal share.

$$mT_1 < \dots < T_m y_t = x_t^\top \beta_r + \varepsilon_t T_{r-1} < t \leq T_r SSR(T_1, \dots, T_m) \\ = \sum_{r=1}^{m+1} \min_{\beta_r} \sum_{t=T_{r-1}+1}^{T_r} (y_t - x_t^\top \beta_r)^2 D_t$$

3.8. Integrated classification rule

Based on previous constructions, a period is classified as robust nearshoring if three conditions are met: the pure geographical effect is negative, ; the change is not explained solely by the disappearance of trade, which requires extensive and intensive margins to be checked; and the queue risk does not increase above an accepted threshold. Expansive reshoring also imposes . Diversification is measured separately by and $.G_i = \sum_k \bar{q}_{ik} \Delta \mu_{ik} < 0 \Delta \logit(R_{ik}) > 0 \Delta S_{ik}^D > 0 \Delta H \Delta E^*$

A synthetic score can be defined only after standardization of the components: , with non-negative weights that add up to one. However, the score does not replace the identification conditions. Two economies can have the same, one through reshoring and the other through nearshoring; Component reporting is mandatory. In a policy issue, weights can be determined by minimizing a losing function or by analyzing the Pareto boundary, not by arbitrary choice. $Z_{it} = w_N z(\Delta N_{it}) + w_R z(\Delta R_{it}) - w_H z(\Delta H_{it}) - w_C z(\Delta C_{it}) Z$

4. Conclusions

The model shows that nearshoring and reshoring are different responses to the same general localization problem. Nearshoring changes the distribution of imports, while reshoring changes the ratio between domestic production destined for the domestic market and imports. The two may be complementary, but there is no identity between them. A firm can shift supply from Asia to Central Europe without increasing production in the country of origin; This is regionalization, not reshoring. Conversely, an increase in domestic production can substitute imports from a nearby country, producing reshoring without a relevant decrease in the average distance of the remaining imports.

Sectoral breakdown is key to interpretation. An aggregate decrease in distance can be generated by the decline of a globalised sector and the growth of a regional sector, even if no sector changes its partners. In the absence of decomposition, that amendment would be erroneously classified as nearshoring. Symmetrically, the rapprochement of partners from several sectors can be masked by the increase in the share of a sector that requires distant

raw materials. The proposed framework avoids both errors and allows variation to be attributed to a geographical and a structural component.

The optimization problem highlights the fact that proximity should not be unconditionally maximized. If nearby suppliers have high costs, limited capacity or strongly correlated shocks, a fully regionalized solution may be dominated by a mixed portfolio. The parameters λ , ρ , and κ provide a direct interpretation of the trade-off: increasing distance aversion moves mass towards proximity, increasing risk aversion favors locations with low covariance, the cost of adjustment produces gradualism, and the preference for diversification limits dependence on a single supplier. In practice, these parameters can be calibrated by the costs of delays, probabilities of interruption and contractual costs. $\lambda\rho\kappa\eta$

Econometric results must be interpreted together with structural indicators. A positive interaction coefficient shows a relative increase in close flows, but does not prove that the average distance decreased if flows to distant partners increased further in other sectors. Similarly, a break in the aggregate series may reflect a price shock, not the relocation of quantities. Therefore, the empirical application must use deflated trade values or quantities, fixed sector-year effects, composition indicators, and robustness analyses at the distance function. $Near \times Post$

The framework also has limits. The economic distance is difficult to observe in full, and the kilometers are only an approximation. Commercial data identifies the place of shipment, not always the origin of the added value. Domestic production and exports can be measured in different classifications, which affects the reshoring share. The risk matrix requires series of interruptions or credible proxies. In addition, the causal relationship between shocks and relocation can be affected by anticipation and selection. These limits do not cancel the model, but establish the conditions under which the results must be interpreted as structural evidence, not as mere correlations. Σ

The paper formulated a unitary mathematical system for the analysis of nearshoring and reshoring. Nearshoring has been defined as a reduction in the geographical timing of commercial distribution and its monotony towards mass transfers to closer partners has been demonstrated. Symmetric decomposition separated the internal geographic effect of sectors from the effect of changing sectoral composition, and extensive and intensive margins identified the mechanism by which reallocation occurs. Reshoring was defined by the share of domestic supply in domestic absorption and logit transformation produced the necessary and sufficient condition that net domestic supply must grow proportionately faster than imports.

The integration of concentration, entropy and risk matrix has demonstrated that closeness is not equivalent to resilience. The convex problem of the supplier structure formalized the compromise between cost, distance, risk, diversification and inertia, and the KKT conditions offered the marginal interpretation of the optimal weights. The PPML specification, the event study and the structural breaks transform the model into an

empirical protocol applicable to bilateral-sectoral flows. The central conclusion is that a reconfiguration can be called nearshoring or reshoring only after separating the volume, composition, geography and origin of the supply. This mathematical discipline allows the comparison of countries and sectors without confusing regionalization, the return of production, diversification and trade contraction.

References

- Anderson, J.E., and Wincoop, V. (2003). Gravity with gravitas: A solution to the border puzzle. *American economic review*, 93(1), 170-192.
- Bai, J., and 3 Perron, P. (2003). Computation and analysis of multiple structural change models. *Journal of applied econometrics*, 18(1), 1-22.
- Boyd, S., and V.L. (2004). *Convex optimization*. 2nd ed. s.l.: Cambridge university press.
- Ciabuschi, F., Lindahl, O., Barbieri, P., and Fratocchi, L. (2019). Manufacturing reshoring A strategy to manage risk and commitment in the logic of the internationalization process model. *European Business Review*, 31(1), 139-159.
- Correia, S., Guimarães, P., and Zylkin, T. (2020). Fast Poisson estimation with high-dimensional fixed effects. *The Stata Journal*, 20(1), 95-115.
- Cuturi, M. (2013). Sinkhorn distances: Lightspeed computation of optimal transport. *Advances in neural information processing systems*, 26(8), 11-37.
- Di Mauro, C.F.L., Orzes, G., and Sartor, M. (2018). Offshoring and backshoring: A multiple case study analysis. *Journal of Purchasing and Supply Management*, 2(2), 108-134.
- Ellram, L.M., Tate, W.L., and Petersen, K.J. (2013). Offshoring and reshoring: an update on the manufacturing location decision. *Journal of Supply Chain Management*, 49(2), 14-22.
- Fratocchi, L. și alții (2014). When manufacturing moves back: Concepts and questions. *Journal of Purchasing and Supply Management*, 20(1), 54-59.
- Gray, J.V., Skowronski, K., Esenduran, G., and Johnny Rungtusanatham, M. (2013). The reshoring phenomenon: what supply chain academics ought to know and should do. *Journal of Supply Chain Management*, 49(2), 27-33.
- Helpman, E., Melitz, M., and Rubinstein, Y. (2008). Estimating trade flows: Trading partners and trading volumes. *The quarterly journal of economics*, 123(2), 441-487.
- Kinkel, S. (2014). Future and impact of backshoring—Some conclusions from 15 years of research on German practices. *Journal of purchasing and Supply Management*, 20(1), 63-65.
- Martínez-Mora, C., and Merino, F. (2014). Offshoring in the Spanish footwear industry: a return journey? *Journal of Purchasing and Supply Management*, 20(4), 225-237.
- McIvor, R., and Bals, L. (2021). A multi-theory framework for understanding the reshoring decision. *International Business Review*, 30(6), 165-187.
- Shannon, C.E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, 27(12), 623-656.

- Silva, J.S., and Tenreyro, S. (2006). The log of gravity. *The Review of Economics and statistics.*, 18(3), 641-658.
- Tate, W.L., Ellram, L.M., Schoenherr, T., and Petersen, K.J. (2014). Global competitive conditions driving the manufacturing location decision.. *Business Horizons.*, 57(3), 381-390.
- Wiesmann, B., Snoci, J.R., Hilletoft, P., and Eriksson, D. (2017). Drivers and barriers to reshoring: a literature review on offshoring in reverse.. *European Business Review.*, 29(1), 15-42.

Analyzing the impact of agricultural growth on inclusive growth in Côte d'Ivoire: key challenges and opportunities

Konan Jules KOUADIO

Shihezi University, China
rulio5884@gmail.com

Koffi Célestin N'GORAN

University Peleforo Gon Coulibaly in Korhogo, Côte d'Ivoire.
nkoffiagos2012@gmail.com

Wei CHANG

Shihezi University, China
shzu_iec@163.com

Abstract. *This study examines how agricultural growth affects inclusive growth in Côte d'Ivoire, focusing on structural transformation and sectoral linkages. Using an Autoregressive Distributed Lag (ARDL) model and Granger causality tests from 1991 to 2021, the findings show that agricultural growth initially negatively impacts non-farm employment but positively influences industrial growth and investment in the long term. The results indicate a unidirectional causality where agricultural growth drives industrial expansion without reverse effects, highlighting weak sectoral linkages. To foster inclusive growth, the study recommends policies enhancing sectoral connectivity through tax incentives, industrial cluster development, and foreign investment promotion.*

Keywords: non-farm employment, agricultural growth, industrial growth, ARDL model.

JEL Classification: O1, O11, O14.

Introduction

Although African's economic growth is mainly driven by the production and export of primary products, the benefits accrue to small groups within the broader economy. The extractive sector, while an important driver of growth, creates few jobs, mainly in urban areas, and is often associated with inequality and social conflict (Ferguson, 2005; Nunnenkamp and Spatz, 2003). In contrast, industry has generally contributed more to GDP growth than agriculture and services, but this unbalanced growth explains the limited progress in poverty reduction despite overall economic growth (Franks, S, 2005). Agriculture is vital to Africa's economy, providing income for 90% of the rural population and accounting for about 20% of total export value (Benedict. S. Kanu et al., 2014). It employs a significant portion of the labor force, with women making up 47% of agricultural workers. As Africa's population is projected to double by 2050, agricultural production must increase substantially to meet growing demands, exacerbated by urbanization and changes in consumption patterns (Thomas and Zuberi, 2012).

The concept of inclusive growth emphasizes expanded access to sustainable socio-economic opportunities for the majority of the population, while protecting vulnerable groups (UNECA and AU, 2012)⁽¹⁾. In Côte d'Ivoire, addressing the challenges of inclusive growth is crucial. Inclusive development requires a coordinated approach that values agriculture while strengthening its linkages with other economic sectors. By improving agricultural productivity and promoting sectoral integration, it is possible to reduce poverty and increase economic opportunities for the majority of the population. In line with the African Development Bank's objectives (African Development Bank, 2023), the government of Cote d'Ivoire has implemented several strategic plans, mainly the National Agricultural Investment Plan (PNIA) and the Ivorian Strategic Development Plan (CSP, 2023-2028) to overcome the challenges facing Côte d'Ivoire's agricultural sector. These plans focus on enhancing the processing of agricultural products, with a target of processing at least 50% of such products. Moreover, this approach seeks to improve the sector's competitiveness, ensure food security, and create jobs, thereby contributing to inclusive growth and structural transformation. This article examines the impact of agricultural growth on inclusive growth in Côte d'Ivoire, focusing on key challenges and opportunities in fostering broader economic inclusivity.

Many studies have demonstrated positive effect of inclusive growth on economic development when sectoral development are taken into account (Lewis, 1954; Fei & Ranis, 1964; Matsuyama, 1992). Mainly, agricultural-led growth theory involves enhancing productivity to catalyze broader economic expansion, improving overall welfare by generating surpluses that invigorate the economy and facilitate job creation across sectors (Gollin 2010; Johnston and Mellor 1961; Kuznets 1961; Kuznets and Murphy, 1966; Ngai and Pissarides, 2007) as well as international organization (Bank, 2007; FAO, 2020). Indeed, agriculture's development is considered as a source of surplus labor and resources for industrialization. Subsequent models emphasized agriculture's provision of food, raw materials, and savings for investment in other sectors (Diao et al., 2018; Gollin, 2010; Johnston and Mellor, 1961; Kuznets, 1961). Recent studies highlight the dynamic

interactions between agriculture and other sectors, demonstrating that agricultural growth stimulates industrialization and structural transformation (Christiaensen, et al., 2006; Christiaensen, et al., 2011; Herrendorf et al., 2014; Pingali 2012; Vollrath, 2011). Agricultural growth fosters inclusive growth by enhancing productivity and incomes within the sector, benefiting the broader economy and creating employment opportunities beyond agriculture (Anderson, 2010; Autor et al., 2003; De Vries et al., 2015; Dollar et al., 2016; Eichengreen & Gupta, 2013; Swinnen, 2018).

Although some countries have struggled to achieve inclusive growth despite notable progress in agricultural development (Dollar, 2004; Dollar & Kraay, 2004), sectoral integration remains crucial for addressing socio-economic challenges. It shapes an economy's capacity to reduce inequality, poverty, and unemployment alongside economic growth (Rodrik et al., 2016). Empirical evidence from countries such as China and several European nations demonstrates that agriculture has played a significant role in poverty alleviation and job creation (Dollar et al., 2016; Loizou et al., 2019; Pingali, 2012; Ravallion & Chen, 2009).

In the case of India, Aggarwal (2018) observed that employment opportunities in high-productivity sectors have not expanded sufficiently due to trade-induced economic specialization and weakened cross-sectoral linkages. This finding underscores the importance of fostering inclusive growth through robust linkages between economic growth and employment generation, as well as between growth and poverty reduction (Aaronson et al., 2014; Sawtelle, 2007; Sodipe & Ogunrinola, 2011). Furthermore, various studies emphasize that integrating agriculture with industry enhances economic profitability, viability, and sustainability (Acemoglu & Guerrieri, 2008; Bustos et al., 2016; Ngai & Pissarides, 2007). For instance, evidence suggests that bidirectional causality between agriculture and industry yields substantial benefits for both sectors (Chaudhuri & Rao, 2004; Henneberry, Khan, & Piewthongngam, 2000). Similar outcomes have been observed in the Brazilian economy (Bustos et al., 2016b). Moreover, industrial development is significantly shaped by the transition from extensive to intensive production systems, which are often driven by innovation and capital deepening (Berthélemy & Söderling, 2001; Chen et al., 2011; Ngai et al., 2019). In this context, technological advancements in agriculture serve as a catalyst for structural transformation and broader economic modernization (Anderson, 2010; Autor et al., 2003; Swinnen, 2018).

In Côte d'Ivoire, agriculture plays a multifaceted role in economic and inclusive growth (World Bank, 2021; FAO, 2020). Studies have highlighted the positive impact of agricultural exports on economic growth (Keho, 2020; Paul-Alfred, 2020) and the importance of agricultural manufacturing, human capital, and foreign direct investment for job creation (Kouakou, 2023). Concerning inclusive growth, several institutional and socio-economic factors influence its dynamism. For instance, research using principal component analysis to calculate an inclusiveness index revealed that household consumption expenditures, bureaucracy, and population growth improve the inclusiveness of Ivorian growth, while GDP per capita and democracy hinder it (Botchuin, 2021; Adeosun et al., 2022).

However, the Ivorian agricultural sector faces significant challenges, including the impacts of climate change, soil degradation, and deforestation (Kouassi et al. 2021). To address these challenges, the government has launched the National Agricultural Investment Plan II (PNIA II), aiming to modernize rural infrastructure and promote resilient agricultural practices (World Bank, 2022; Farmonaut, 2024). Additionally, initiatives to strengthen agricultural market regulation and encourage crop diversification are underway to reduce dependence on cocoa and improve sector resilience against climate-related risks.

This research aims to fill the research gap by examining the impact of agricultural growth on non-farm employment and structural transformation using an ARDL model and Granger causality tests. The main hypothesis is that agricultural growth leads to job creation in non-farm sectors, contributing to inclusive growth. The study will explore the characteristics of economic growth in Côte d'Ivoire and assess whether agricultural productivity can contribute to inclusive growth. Two primary questions will be addressed: What are the characteristics of economic growth in Côte d'Ivoire? Can agricultural productivity contribute to inclusive growth in Côte d'Ivoire?

The study is organized into sections that present the introduction and literature review (1), followed by conceptual framework in section (2), The data description and descriptive statistics in section (3), the results and discussion in the section (4), and conclude with recommendations in the section (5). By exploring these questions, the research aims to provide insights into how agricultural growth can be leveraged to enhance inclusive growth and structural transformation in Côte d'Ivoire.

Conceptual framework

This section presents theoretical and empirical framework, the description of the variables and the estimation methodology.

Agricultural Growth and Inclusive Development in Côte d'Ivoire

Given that the core theme of this paper is agricultural growth and inclusive development, this study proposes an operational definition of inclusive growth within the context of Côte d'Ivoire. For growth to be truly inclusive in the Ivorian context, agricultural expansion must generate economy-wide benefits specifically through employment creation and income gains while emphasizing the importance of intersectoral linkages. This resonates with literature emphasizing that inclusive growth should not merely focus on increasing GDP growth rates but rather on improving productivity and the living conditions of the poor through a more equitable distribution of the gains from economic growth (Fofana et al., 2022).

As agriculture in Côte d'Ivoire is often constrained by the availability of arable land and employs roughly 45% of the total labor force (FAO, 2023), the degree of bidirectional integration between the agricultural sector and other economic sectors is critical in

determining the success or failure of the country's structural transformation policies. Yet, this dimension is frequently overlooked in macroeconomic assessments of national performance toward emergence.

The underlying assumption of this framework is that growth is inclusive when it allows the majority of society to participate in, benefit from, and contribute to the growth process. Agriculture can serve as a critical lever for stimulating employment, raising incomes, and improving living standards—especially when previously marginalized populations are granted access to decent and productive employment. The long-term effects of inclusive growth include reduced unemployment, lower dependency on food aid and public subsidies, and increased government revenues. Higher employment and income levels are often associated with lower crime rates and greater political stability key conditions for attracting foreign direct investment (Andriano, 2024).

This paper thus argues that adopting an inclusive approach to agricultural development within a broader policy reform agenda is necessary to ensure the equitable distribution of economic benefits and employment opportunities across both rural and urban populations. Such an approach ensures that the progress made in agriculture translates into tangible social and economic outcomes for the wider population.

Despite Côte d'Ivoire's rich natural resource base, the agricultural sector faces major challenges in meeting the country's development expectations. Agricultural growth has been primarily driven by the expansion of cultivated land, rather than by improvements in productivity a path that is unsustainable in the long term (Fofana et al., 2022). Moreover, per capita agricultural GDP growth has remained limited, partly due to the parallel expansion of the agricultural labor force. Compounding this, the FAO has warned that climate change will continue to put substantial pressure on renewable water resources needed for irrigation a stress that is expected to intensify in the coming decades (FAO, 2023; Watt & Chamberlain, 2011). This scenario poses significant risks to agricultural output in Côte d'Ivoire between 2080 and 2100, with potentially large production losses.

These dynamics underscore the urgency of shifting toward productivity-led agricultural strategies that are environmentally sustainable and intersectorally integrated. Only through such approaches can agricultural growth serve as a driver of inclusive development and lead to more resilient and equitable economic outcomes.

Theoretical and empirical modeling

Theoretical modeling

The relationship between agricultural growth and non-agricultural employment is essential for inclusive growth and structural transformation in developing economies. To achieve this objective, we construct a model based on the model of economic growth developed by (Barro 1990; Lucas Jr 1988; Mankiw, Romer, and Weil 1992; Romer 1990) which is an endogenous growth model.

The theoretical functional form of the endogenous growth model is given by:

$$Y = f(K, L, X) \quad (1)$$

Where Y represents gross domestic product, K is capital, L is labor, and X represents other exogenous variables.

From this, we specify the theoretical ARDL model

$$\Delta y_t = \alpha + \sum_{i=1}^p \beta_i \Delta y_{t-i} + \sum_{i=0}^q \gamma_i \Delta X_{t-i} + \delta_1 \Delta y_{t-1} + \delta_2 \Delta x_{t-1} + \varepsilon_t \quad (2)$$

Where Y is the dependent variable, X are the exogenous variables, Δ is the first difference operator, and ε is the error term.

The Error Correction Model (ECM) is obtained by:

$$\Delta y_t = \alpha + \sum_{i=1}^p \beta_i \Delta y_{t-i} + \sum_{i=0}^q \gamma_i \Delta X_{t-i} + \phi ECT_{t-1} + \varepsilon_t \quad (3)$$

Where ECT is the error correction term representing the long-run relationship and ϕ is the speed of adjustment coefficient.

To analyze the dynamic causality between variables, we apply the Granger causality test:

$$y_t = \sum_{i=1}^p \alpha_i \Delta y_{t-i} + \sum_{i=0}^q \beta_i X_{t-i} + \varepsilon_t \quad (4)$$

Where p and q are the maximum lags considered for Y and X respectively. It is a statistical tool to test whether one-time series can predict another time series. The concept, developed by Clive Granger in 1969, is based on the idea that if the prediction of a time series can be improved by incorporating the past values of another time series, then is said to Granger-cause.

Empirical modeling

The autoregressive distributed lag (ARDL) estimation method is a robust statistical technique for analyzing long-term relationships between variables in time series data, accommodating variables that are stationary at level $I(0)$, at the first difference $I(1)$, or a mix of both. The process begins with the collection of relevant data, followed by a descriptive analysis to assess the stationarity, central tendencies, and variability of the variables. Subsequently, the ARDL model is estimated using standard linear regression, incorporating appropriate lags for both dependent and explanatory variables to capture the dynamic relationships. The model's robustness is evaluated through several residual diagnostic tests: the CUSUM test for parameter stability, the Breusch-Godfrey LM test for autocorrelation, the White test for heteroscedasticity, the Shapiro-Wilk test for normality, and the variance inflation factor (VIF) test for multicollinearity. These tests ensure the

reliability and validity of the model, confirming that it provides consistent and unbiased estimates. The following empirical ARDL model is used in this study:

$$\begin{aligned} \Delta NFEMP_t = & \alpha_0 + \sum_{i=1}^p \alpha_1 \Delta NFEMP_{t-i} + \sum_{i=0}^{q_1} \beta_1 \Delta AGRV_{t-i} + \sum_{i=0}^{q_1} \beta_2 \Delta INDV_{t-i} \\ & + \sum_{i=0}^{q_1} \beta_3 \Delta SAV_{t-i} + \sum_{i=0}^{q_1} \beta_4 \Delta GFCF_{t-i} + \sum_{i=0}^{q_1} \beta_5 \Delta EXR_{t-i} + \gamma_1 \Delta NFEMP_{t-i} \\ & + \gamma_2 \Delta AGRV_{t-i} + \gamma_3 \Delta INDV_{t-i} + \gamma_4 \Delta SAV_{t-i} + \gamma_5 \Delta GFCF_{t-i} \\ & + \gamma_6 \Delta EXR_{t-i} + \varepsilon_t \end{aligned} \quad (5)$$

The Error Correction Model (ECM) is obtained by the following equation:

$$\begin{aligned} \Delta NFEMP_t = & \alpha_0 + \sum_{i=1}^p \alpha_1 \Delta NFEMP_{t-i} + \sum_{i=0}^{q_1} \beta_1 \Delta AGRV_{t-i} + \sum_{i=0}^{q_1} \beta_2 \Delta INDV_{t-i} \\ & + \sum_{i=0}^{q_1} \beta_3 \Delta SAV_{t-i} + \sum_{i=0}^{q_1} \beta_4 \Delta GFCF_{t-i} + \sum_{i=0}^{q_1} \beta_5 \Delta EXR_{t-i} + \phi ECM_{t-1} \\ & + \varepsilon_t \end{aligned} \quad (6)$$

Where:

$NFEMP$ is the Non-Farm employment, $AGRV$ is Agricultural Value Added, $INDV$ is Industrial Value Added, SAV is Service Value Added, $GFCF$ is Gross Fixed Capital Formation, EXR is Exchange rate, Δ denotes the first difference operator, α_0 is the constant term, $\alpha_i, \beta_i, \gamma_i$ are the coefficients, ε_t is the error term, ECM_{t-1} is the error correction term representing the long-term relationship; ϕ is the speed of adjustment coefficient.

Data description and descriptive statistics

Data description

The empirical model utilizes data from 1991-2021, primarily from the World Bank except Agricultural Added Value ($AGRV$) drawn from FAO database. Key variables include: Non-Farm Employment Rate ($NFEMP$) obtained by the sum of industry and service employment as % of total employment, Agricultural Added Value ($AGRV$), Industrial Added Value ($INDV$), Service Added Value (SAV) representing respectively the value added by agriculture, value added by industry, and value added by services as % of GDP. Gross Fixed Capital Formation ($GFCF$) indicating the net increase in fixed assets and investment as % of GDP and the Exchange Rate (EXR) which reflects changes in import/export costs for the country. The GDP is desegregated into Agricultural Added Value ($AGRV$) representing the value added by crop production, livestock, forestry, and fishing, Industrial Added Value ($INDV$) including activities such as mining, manufacturing, construction, and utilities, while Service Added Value (SAV) covers a wide range of activities including retail, transport, finance, real estate, and public administration.

To analyse the relationship between agricultural productivity growth and inclusive growth in Côte d'Ivoire, it is essential to consider a sufficiently long period to capture the effects of economic reforms and assess long-term impacts. The economic reforms initiated in the late 1970s, following the agricultural commodities crisis, led to the widespread adoption of structural adjustment programs across developing economies in the early 1990s. In Côte d'Ivoire, these reforms profoundly transformed the economic system, with significant implications for the agricultural sector and inclusive growth. To accurately evaluate these impacts, it is necessary to encompass a sufficiently extended timeframe to reflect economic adjustments and variations in agricultural productivity. Thus, the study covers a thirty-year period, allowing for the observation of long-term trends beyond the immediate effects of the reforms. This period also ensures that the available data are robust enough to adhere to the normal distribution law, which is crucial for reliable statistical analysis. This approach ensures a comprehensive and well-founded assessment of the reforms and their effects on agricultural productivity and inclusive growth, accounting for both immediate impacts and long-term changes.

Descriptive statistics

Table 1 presents a summary of the descriptive statistics. Descriptive statistics show the means are higher than standard errors, indicating the sample means are likely good representations of the population means. Additionally, the fact that the standard errors are less than the means suggests that the data points are clustered relatively closely around the mean, indicating a lower degree of variability in the data, which is beneficial for our model.

Table 1: *Descriptive statistics*

Variables	Mean	Std.error
Non-farm employment(% Total employment)	50.73	0.69
Agriculture added value(%GDP)	18.88	0.9
Industry added value(%GDP)	18.11	0.57
Service added value (%GDP)	42.65	1.45
Gross Capital Formation	16.78	0.83
Exchange Rate (%)	96.54	1.48

Source: Authors

Table 1: *Unit root test ADF*

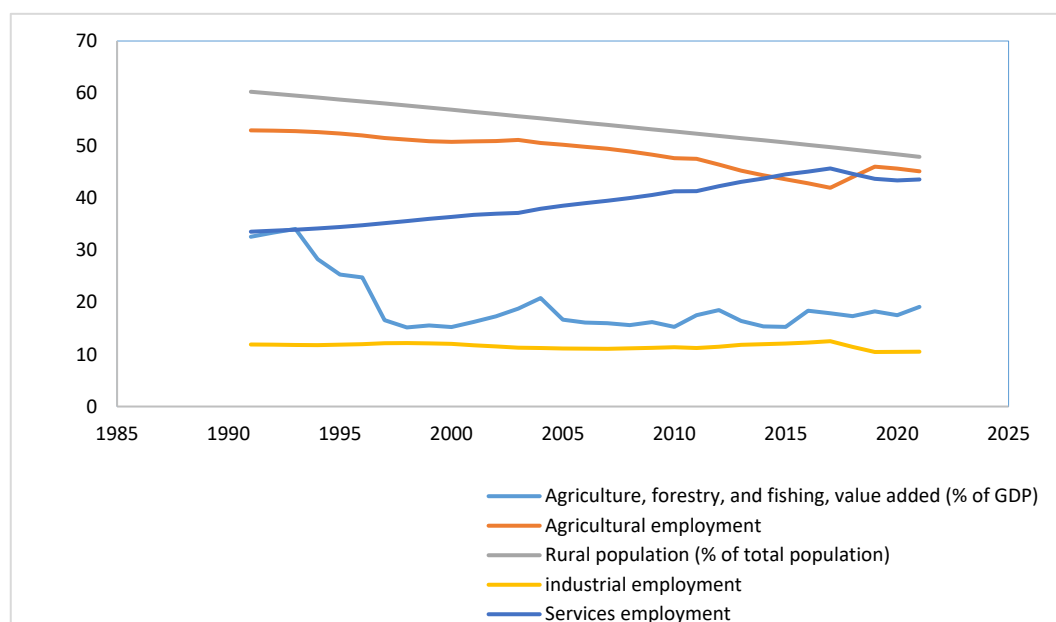
Variables	Test Type	Level/First Difference	Test Statistic
Non-farm employment	ADF	I(0)	-2.69
Non-farm employment	ADF	I(1)	-3.23***
Agricultural added value	ADF	I(0)	-4.13***
Agricultural added value	ADF	I(1)	-3.69***
Industry added value	ADF	I(0)	-3.09
Industry added value	ADF	I(1)	-3.83***
Services added value	ADF	I(0)	-1.91
Services added value	ADF	I(1)	-4.28***
Gross fixed capital formation	ADF	I(0)	-2.45
Gross fixed capital formation	ADF	I(1)	-4.89***
Exchange rate	ADF	I(0)	-1.63
Exchange rate	ADF	I(1)	-5.40***
H ₀ : Null hypothesis: The series has a unit root/ H ₁ :No H ₀ :			

Sources: Authors

The Augmented Dickey-Fuller (ADF) reported in the **table 2** test reveals that most variables are non-stationary at levels but stationary after first differencing, suggesting regressions at level form may be spurious, with the exception of agricultural added value, which is stationary at level.

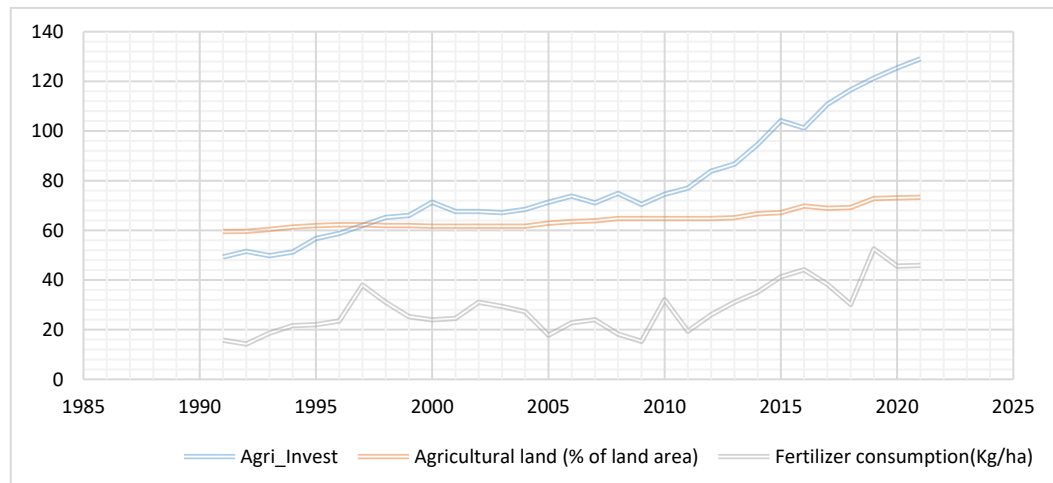
Figure 1 shows a significant decline in agriculture's GDP and employment share alongside reduced rural population, while industry and services employment grew substantially. This reflects a shift towards more productive sectors, increased resilience, and potential for sustainable growth. This economic characteristic reflects a shift towards more productive sectors. This transition underscores urbanization trends and highlights the economy's move towards diversification, improved productivity, and higher income levels. These changes suggest increased economic resilience and potential for sustainable growth, although they necessitate supportive policies to ensure that the benefits of this transformation are inclusively shared across the population.

Figure 1: Evolution of employment and agricultural value added by sector



Source: Authors

Figure 2 shows rising trends of agricultural investment, land utilization, and fertilizer use, indicating structural change in the sector. Increased investment stimulates productivity, efficiency and economic growth through advanced technologies which leads to higher output and job creation, particularly in rural areas impacting poverty alleviation. Efficient land utilization and fertilizer use maximize crop yields and improve food security by increasing food availability and stabilizing prices. Furthermore, these improvements facilitate structural transformation by reallocating labor to higher-productivity sectors, promoting economic diversification, and adding value through agro-processing.

Figure 2: Trends in the use of agricultural inputs

Source: Authors

Results and discussions

This section presents the robustness tests which validate the ARDL model's reliability

Robustness tests associated with the ARDL model results

Table 3 presents the Variance Inflation Factor (VIF) test results. To ensure the variables reliability, we tested for multicollinearity using VIF, as high multicollinearity can distort estimated coefficients and reduce model precision. We specified the ARDL model and selected optimal lags for each variable based on the Akaike Information Criterion (AIC). The VIF test, is conducted on the explanatory variables and their lagged terms and is interpreted as follows: $VIF < 5$ indicates no serious multicollinearity, $5 \leq VIF < 10$ indicates moderate multicollinearity, and $VIF \geq 10$ indicates high multicollinearity (corrective actions needed). The Variance Inflation Factor (VIF) test indicates no serious multicollinearity.

Table 2: Variance Inflation Factor (VIF) test

Variable	Lags	VIF	1/VIF
Ln(Services added value)	0	8.42	0.118
Ln(Agriculture added value)	1	8.23	0.121
Ln(Gross fixe capital formation)	0	7.09	0.141
Ln(Industry added value)	1	6.88	0.145
Ln(Industry added value)	0	6.68	0.149
Ln(Gross fixe capital)	2	6.12	0.163
Ln(Exchange rate)	1	5.78	0.173
Ln(Exchange rate)	0	5.6	0.178
Ln(Services added value)	1	5.12	0.195
Ln(Agricultural added value)	0	4.65	0.214
Ln(Gross fixe capital formation)	1	4.65	0.214
Mean		VIF	6.29

Sources: Authors

Table 4 presents the cointegration test results. It shows that the Fisher statistic (5.5) exceeds the upper bounds, indicating a long-run relationship between the variables. Consequently, both short-run and long-run equations will be estimated.

Table 3: Bound test of cointegration

Level of significance (α %)	Lower value	Upper value	Hypothesis
5%	2.62	3.79	H0:No level relationship
10%	2.26	3.35	
Accept if F<critical calculate for I (0) regressors, reject if F>to critical value Computed F-statistic: 5.5			

Source: Authors; * denotes significance at 5% and 10% level

Table 5 reports the results of the robustness tests. It emerges from the results the absence of autocorrelation, of heteroscedasticity and the errors are normally distributed. The validation of an ARDL model requires the homoscedasticity of residuals, a normal distributed error term and no serial correlation on the residuals. The results in table 4 demonstrate the model's robustness.

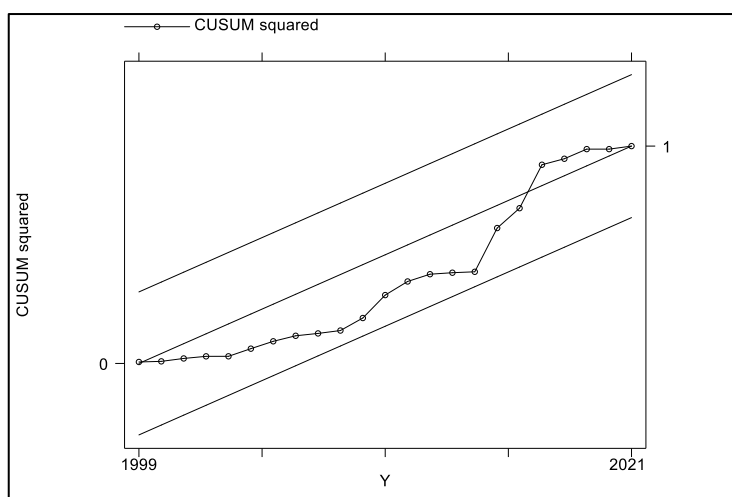
Table 4: Results of robustness test

Tests	Hypothesis	P>chi2	Decision
Breusch–Godfrey LM test	H0: No serial correlation/H1: non H0	0.412	Do not reject H0
White's test	H0: Homoscedasticity/H1: non H0	0.412	Do not reject H0
Ramsey RESET test (normality)	H0: Model has no omitted variables/H1: non H0	0.014	Do not reject H0

Source: Authors

Figure 3 record residuals stability test. After estimation, we verify the stability of the model's coefficients over time, at the end of the forecast. We use *Brown-Durbin-Ewans (1975)* tests. This test allows us to detect structural instabilities. Globally there is instability if the curve crosses the corridor. It appears that the model does not have a structural break. Therefore, the model is stable and its forecasts and inferences are robust.

Figure 3: CUSUM stability test



Source: Authors

Granger causality test results

Table 6 reports the Granger causality test results. According to *Granger (1988)*, in the presence of cointegration, there must be at least one direction of causality: unidirectional or bidirectional. There are various methods for lag length selection in recent literature, including the Akaike Information Criterion (AIC), the Schwartz Information Criterion (SIC), and *Hsiao (1979)*'s sequential procedure. Our results reveal significant causal relationships that underscore the pivotal role of agricultural performance in influencing broader economic variables. The analysis indicates that agricultural added value is the Granger cause of non-farm employment, suggesting that fluctuations in agricultural productivity substantially affect employment levels outside the agricultural sector. This result highlights the importance of agricultural sector in driving labor market dynamics, influencing non-farm job creation and supported the theory according to which agriculture-led growth can stimulate the broader economy by increasing rural incomes, which in turn raises demand for goods and services produced in other sectors (*Johnston and Mellor, 1961*).

Table 5: Matrix of Granger causality test

Null hypothesis : Independent variable does not cause the dependent variable				
Dependent variable (1)	Independent variables(2)	F-Statistic	Lags	Conclusion
D_InNon-farm employment	D.InAgricultural added value	6.37*** (0.095)	3	Reject
D_InNon-farm employment	D.InIndustry added value	15.28*** (0.002)	3	Reject
D_InNon-farm employment	D_InServices added value	20.19*** (0.00)	3	Reject
D_InNon-farm employment	D_InGross fixe capital formation	15.42*** (0.001)	3	Reject
D_InNon-farm employment	Exchange rate	16.51*** (0.001)	3	Reject
D_InAgricultural added value	D.InIndustry added value	10.2 (0.2)	3	Accept
D_InAgricultural added value	D_InServices added value	12.02*** (0.007)	3	Reject
D_InIndustry added value	D.InAgricultural added value	23.02*** (0.00)	3	Reject
D_InIndustry added value	D_InServices added value	46.30*** (0.00)	3	Reject
D_InServices added value	D.InIndustry added value	11.51*** (0.009)	3	Reject
D_InGross fixe capital formation	D.InAgricultural added value	8.02*** (0.04)	3	Reject

Notes: P-value in parentheses; Standard errors in parentheses * $p < 0.01$, ** $p < 0.05$, *** $p < 0.1$ significance level.

Sources: Authors

Moreover, the findings stipulate that agricultural growth is the granger causes of industrial and services expansion, and improve gross fixed capital formation. It means that agricultural growth stimulates demand for industry, services and improves investment through saving. In other word, agricultural growth assumes its role of forward and backward linkages by spurring growth in other sectors. This finding is consistent with *Timmer's (2005)* study, who argued that agricultural productivity leads to higher rates of investment in physical capital, thereby promoting overall economic development. However, the industrial growth does not affect agricultural sector highlighting a key challenge identified in many developing countries studies in the economic literature. This lack of integration limits the potential for mutually reinforcing growth between these sectors, as found by *Gollin et al. (2002)*. These results show that agricultural growth plays a crucial role in economic development by influencing non-farm employment, stimulating industrial and service sector growth, and improving investment in capital formation in Cote d'Ivoire.

ARDL estimation report

The ARDL model results are recorded in **table 7**, providing an insight of the short-run and long-run relationships between non-farm employment and the explanatory variables. The adjustment term with a negative and significant (at 1% level) coefficient of about 0.288, which intuitively means that a possible shock on non-farm employment is corrected at the convergence speed of 28.8% in the area of study. Moreover, the coefficient of the lagged non-farm employment is positive and significant, indicating about 59.3% stability of non-farm employment over the study period. Moreover, agricultural and industrial growths have a negative and significant effect on non-farm employment, with respective coefficients of 0.058 and 0.0567. This suggests that a 1% increase in these sectors will decrease the non-farm employment of about 5.8% and 5.67%, respectively and it is due to the fact that agricultural sector in Cote d'Ivoire employs more than 46% of total worker force and is still the leading sector in terms of employment according to FAO. Furthermore, the development of agricultural sector causes a labor surplus oriented toward nonagricultural sector. As shown in Table 5, industry growth does not cause agricultural sector growth, thereby limiting agriculture's role in job creation within the value chain. These results align with structural transformation literature *Timmer* (1988), highlighting labor shifts from non-farm sectors and short-term labor market inefficiencies in responding to sector-specific growth. Gross capital formation (proxy to investment) and exchange rates negatively impact non-farm employment, indicating that capital-intensive growth can initially reduce labor demand, similar to trends in automation. The exchange rate effect aligns with open economy macroeconomics, where currency appreciation reduces export competitiveness and employment in export-driven sectors (*Freund and Pierola, 2012*).

Table 6: ARDL estimation result

Variables	Coefficients
Short-run equation	
LD. ln (non-farm employment)	0.593***(-0.127)
D. ln (agricultural added value)	-0.0580**(-0.019)
D. ln (industry added value)	-0.0567***(-0.022)
D. ln (service added value)	0.0695***(-0.016)
D. ln (gross capital formation)	-0.0708***(-0.017)
LD. ln (gross capital formation)	-0.0490***(-0.013)
D. ln (exchange rate)	-0.201***(-0.045)
Long-run equation	
ADJ	-0.288***(-0.075)
Ln (agricultural added value)	-0.065(-0.075)
Ln (industry added value)	0.357**(-0.090)
Ln (service added value)	-0.203***(-0.086)
Ln (gross capital formation)	0.211**(-0.044)
Ln (exchange rate)	0.360**(-0.103)
Constant	0.362*(-0.175)
Observations	29
R-squared	0.858
Adj R-squared	0.735
Log likelihood	112.638

Standard errors in parentheses * p<0.01, ** p<0.05, *** p<0.1

Sources: Authors

Conversely, service sector growth positively impacts non-farm employment, indicating that it boosts job creation in the early stages of economic development (*Eichengreen and Gupta, 2013b*). This is supported by *De Vries et al. (2015)*, who found that the service sector is increasingly driving economic growth and job creation in developing countries.

Agricultural growth does not directly impact non-farm employment in the long run; however, industrial growth and gross capital formation positively influence it. Granger causality and ARDL results reveal that agricultural growth indirectly boosts non-farm employment by supporting industrial growth and generating investment income, as noted by (*Kuznets, 1961; Lucas, 1953*). A 1% increase in the industrial sector and gross capital formation leads to a 35.7% and 21.1% rise in non-farm employment, respectively. These findings align with *McMillan, et al. (2017)*, emphasizing industrialization as crucial for higher productivity and job creation in developing countries. Conversely, a negative and significant effect of services growth on non-farm employment is found in the long-run. It shows that the services sector is more labour-extensive. The exchange rate has a positive and significant effect on non-farm employment, suggesting that a depreciation of the local currency enhances non-farm employment by boosting export competitiveness as also founded by (*Rodrik, 2008*). In the long run, the exchange rate positively affects non-farm employment, suggesting that currency depreciation boosts export competitiveness and thus increases non-farm employment, consistent with *Rodrik (2008)*. Conversely, service sector growth negatively impacts non-farm employment, indicating its labor-extensive nature.

Conclusion and recommendations

This study analysed the effect of agricultural growth on inclusive growth through the dynamic relationship between agricultural growth and non-farm employment in Côte d'Ivoire. It applied a combined approach of an ARDL model and Granger causality test on data covering the period 1991 to 2021 and provided by World Bank and FAO. Preliminary results show that non-farm employment adjusts to shocks at a 28.8% speed, with 59.3% stability shown by the lagged term in the short run. Key findings stipulate that agricultural growth negatively impacts non-farm employment in the short-run due to labor shifts and inefficient labor markets. It emerges of the both approaches alignment that agricultural growth positively affects industrial expansion, investment, and exchange rate conditions. Consequently, agricultural growth indirectly supports non-farm employment through industrial growth and capital formation. However, the lack of bidirectional causality between agricultural and industrial growth, limits agriculture's role in supporting non-farm job creation.

To achieve significant structural transformation for inclusive growth, the study recommends policies that strengthen the agriculture-industry connection. This includes targeted tax incentives, industrial clusters, and training programs to build skills and support innovation. Encouraging businesses to pursue international opportunities and cross-sector

partnerships will further enhance these efforts. Moreover, future research should further explore the specific contributions of different agricultural branches to inclusive growth and the effectiveness of policy interventions aimed at enhancing these linkages.

Note

⁽¹⁾ UNECA et AU : www.uneca.org

References

- Aaronson, D., Hu, L., Seifoddini, A., and Sullivan, D.G. (2014). Declining Labor Force Participation and Its Implications for Unemployment and Employment Growth. *Economic Perspectives* 38(4).
- Acemoglu, D., and Guerrieri, V. 2008. Capital Deepening and Nonbalanced Economic Growth. *Journal of Political Economy* 116(3):467–98.
- Adeosun, O.A., and Tabash, M.I. (2022). Pro-poor and inclusive growth in West Africa. *African Journal of Economic and Management Studies*, 13(1), 105-135.
- Aggarwal, A. (2018). Economic Growth, Structural Change and Productive Employment Linkages in India: Did Market Transition Matter? *South Asia Economic Journal* 19(1):64–85.
- Anderson, K. (2010). *The Political Economy of Agricultural Price Distortions*. Cambridge University Press.
- Andriano, L. (2024). Agricultural droughts and under-five mortality in Côte d'Ivoire: Differential impacts across social and demographic groups*. *Population and Environment*. <<https://doi.org/10.1007/s11111-024-00469-0>>
- Autor, D.H., Levy, F., and Murnane, R.J. (2003). The Skill Content of Recent Technological Change: An Empirical Exploration. *The Quarterly Journal of Economics* 118(4):1279–1333.
- Bank (2007). *World Development Report 2008: Agriculture for Development*. World Bank Publications.
- Barro, R.J. (1990). Government Spending in a Simple Model of Endogeneous Growth. *Journal of Political Economy* 98(5, Part 2):S103–25. doi: 10.1086/261726.
- Berthélemy, J.-C., and Söderling, L. (2001). The Role of Capital Accumulation, Adjustment and Structural Change for Economic Take-off: Empirical Evidence from African Growth Episodes. *World Development* 29(2):323–43.
- Botchuin, W.C. (2021). Institutional Factors of Inclusive Growth: Evidence from Cte Divoire. *Journal of Economics and International Finance* 13(1):13–23.
- Bustos, P., Caprettini, B., and Ponticelli, J. (2016a). Agricultural Productivity and Structural Transformation: Evidence from Brazil. *American Economic Review* 106(6):1320–65.
- Chaudhuri, K., and Rao, R.K. (2004). Output Fluctuations in Indian Agriculture and Industry: A Reexamination. *Journal of Policy Modeling* 26(2):223–37.
- Chen, S., Jefferson, G.H., and Zhang, J. (2011). Structural Change, Productivity Growth and Industrial Transformation in China. *China Economic Review* 22(1):133–50.

- Christiaensen, L., Demery, L., and Kuhl, J. (2011). The (Evolving) Role of Agriculture in Poverty Reduction—An Empirical Perspective. *Journal of Development Economics* 96(2):239–54.
- Christiaensen, L.J., Demery, L., and Khl, J. (2006). *The Role of Agriculture in Poverty Reduction: An Empirical Perspective*. Vol. 4013. World Bank Publications.
- De Vries, G., Timmer, M., and De Vries, K. (2015). Structural Transformation in Africa: Static Gains, Dynamic Losses. *The Journal of Development Studies* 51(6):674–88. <doi: 10.1080/00220388.2014.997222>
- Diao, X., McMillan, M., and Wangwe, S. (2018). Agricultural Labour Productivity and Industrialisation: Lessons for Africa. *Journal of African Economies* 27(1):28–65.
- Dollar, D., and Kraay, A. (2004). Trade, Growth, and Poverty. *The Economic Journal* 114(493):F22–49.
- Eichengreen, B., and Gupta, P. (2013a). The Real Exchange Rate and Export Growth: Are Services Different? Bank of Korea WP 17.
- FAO (2020). The State of Food Security and Nutrition in the World. Disponible en ligne: <www.fao.org>
- Farmonaut (2024). Sustainable Agricultural Transformation in Côte d'Ivoire: Innovations and Opportunities for Green Economic Growth. Available at <<https://farmonaut.com/africa/transformation-agricole-durable-en-cote-divoire-innovations-et-opportunités-pour-une-croissance-economique-verte/>>
- Fei, J.C.H., and Ranis, G. (1964). Capital-Labor Ratios in Theory and in History: Reply. *The American Economic Review* 1063–69.
- Ferguson, J. (2005). Seeing like an oil company: space, security, and global capital in neoliberal Africa. *American anthropologist*, 107(3), 377-382.
- Fofana, I. et al. (2022). *Agricultural development promises more growth and less poverty in Africa. *Development Policy Review*, 40(6), e12669. <<https://doi.org/10.1111/dpr.12669>>
- Franks, S. (2005). Our common interest: Report of the commission for Africa. Political Quarterly-London then Oxford-Macmillan then Blackwell-, 76(3), 446.
- Freund, C., and Pierola, M.D. (2012). Export Surges. *Journal of Development Economics* 97(2):387–95.
- Gollin, D. (2010). Agricultural Productivity and Economic Growth. *Handbook of Agricultural Economics* 4:3825–66.
- Gollin, D., Parente, S., and Rogerson, R. (2002). The Role of Agriculture in Development. *American Economic Review* 92(2):160–64.
- Granger, C.W.J. (1988). Causality, Cointegration, and Control. *Journal of Economic Dynamics and Control* 12(2–3):551–59.
- Henneberry, S.R., Khan, M.E., and Piewthongngam, K. (2000). An Analysis of Industrial–Agricultural Interactions: A Case Study in Pakistan. *Agricultural Economics* 22(1):17–27. doi: 10.1111/j.1574-0862.2000.tb00002.x.
- Herrendorf, B., Rogerson, R., and Valentinyi, A. (2014). Growth and Structural Transformation. *Handbook of Economic Growth* 2:855–941.
- Hsiao, C. (1979). Autoregressive Modeling of Canadian Money and Income Data. *Journal of the American Statistical Association* 74(367):553–60.
- Johnston, B.F., and Mellor, J.W. (1961). The Role of Agriculture in Economic Development. *The American Economic Review* 51(4):566–93.

- Kouakou Kouakou, P.A. (2020). Agricultural Productions and Its Implications on Economic Growth in Ivory Coast: The Use of the Econometric Approach. *Applied Economics and Finance* 7(4):126.
- Kouakou Kouakou, P.-A. (2023). Impact of Local Processing of Agricultural Raw Materials on Job Creation in the West African Monetary and Economic Union. *Agricultural and Resource Economics: International Scientific E-Journal* 9(3):250–65.
- Kuznets, S. (1961). Economic Growth and the Contribution of Agriculture: Notes on Measurement.
- Lewis, W.A. (1954). Economic Development with Unlimited Supplies of Labour.
- Loizou, E., Karelakis, C., Galanopoulos, K., and Mattas, K. (2019). The Role of Agriculture as a Development Tool for a Regional Economy. *Agricultural Systems* 173:482–90.
- Lucas Jr, R.E. (1988). On the Mechanics of Economic Development. *Journal of Monetary Economics* 22(1):3–42.
- Mankiw, N.G., Romer, D., and Weil, D.N. (1992). A Contribution to the Empirics of Economic Growth. *The Quarterly Journal of Economics* 107(2):407–37.
- Matsuyama, K. (1992). Agricultural Productivity, Comparative Advantage, and Economic Growth. *Journal of Economic Theory* 58(2):317–34.
- McMillan, M., Rodrik, D., and Sepúlveda, C. (2017). Structural Change, Fundamentals and Growth: A Framework and Case Studies. National Bureau of Economic Research.
- Ngai, L.R., and Pissarides, C.A. (2007). Structural Change in a Multisector Model of Growth. *American Economic Review* 97(1):429–43.
- Nunnenkamp, P., and Spatz, J. (2003). Foreign direct investment and economic growth in developing countries: how relevant are host-country and industry characteristics? (No. 1176). Kiel working paper.
- Pingali, P.L. (2012). Green Revolution: Impacts, Limits, and the Path Ahead. *Proceedings of the National Academy of Sciences* 109(31):12302–8. doi: 10.1073/pnas.0912953109.
- Ravallion, M., and Chen, S. (2009). “China’s (Uneven) Progress against Poverty”. Pp. 65–111 in *Governing rapid growth in China*. Routledge.
- Rodrik, D. (2008). The Real Exchange Rate and Economic Growth. *Brookings Papers on Economic Activity* 2008(2):365–412.
- Rodrik, D., McMillan, M., and Sepúlveda, C. (2016). Structural Change, Fundamentals, and Growth. *Structural Change, Fundamentals, and Growth*. Washington DC: IFPRI.
- Romer, P.M. (1990). Endogenous Technological Change. *Journal of Political Economy* 98(5, Part 2):S71–102. doi: 10.1086/261725.
- Sawtelle, B. (2007). Analyzing the Link Between Real GDP and Employment: An Industry Sector Approach: Employment Responsiveness to Long-Term Real GDP Expansion Differs Markedly Across US Industry Sectors. *Business Economics* 42:46–54.
- Sodipe, O.A., and Ogunrinola, O.I. (2011). Employment and Economic Growth Nexus in Nigeria. *International Journal of Business and Social Science* 2(11):232–39.
- Swinnen, J. (2018). “The Political Economy of Agricultural and Food Policies”. Pp. 381–98 in *The Routledge Handbook of Agricultural Economics*. Routledge.
- Timmer, C.P. (1988). The Agricultural Transformation. *Handbook of Development Economics* 1:275–331.
- Timmer, P. (2005). *Agriculture and Pro-Poor Growth: An Asian Perspective*. Available at SSRN 1114155.

- Vollrath, D. (2011). The Agricultural Basis of Comparative Development. *Journal of Economic Growth* 16:343–70.
- Watt, S., and Chamberlain, J. (2011). Water, climate change, and maternal and newborn health. Current Opinion in *Environmental Sustainability*, 3(6), 491–496. <<https://doi.org/10.1016/j.cosust.2011.10.001>>
- World Bank (2022). Increased Productivity Can Transform Growth Prospects in Côte d'Ivoire. Available at <<https://blogs.worldbank.org/fr/africacan/une-productivite-accrue-peut-transformer-les-perspectives-de-croissance-en-cote-divoire>>
- Zuberi, T., and Thomas, K.J. (2012). Demographic projections, the environment and food security in sub-Saharan Africa. Regional bureau for Africa, United Nations working group, 22.

Credibility erosion and exchange rate pass-through in Türkiye

Umit BULUT

Kirsehir Ahi Evran University, Türkiye
ubulut@ahievran.edu.tr

Okan DADALI

Kirsehir Ahi Evran University, Türkiye
dadaliokan2000@gmail.com

Azra Asel BAYRAK

Kirsehir Ahi Evran University, Türkiye
bayrak.azraasel@ogr.ahievran.edu.tr

Abstract. *This study aims to examine the impact of the erosion of the Central Bank of the Republic of Türkiye (CBRT)'s credibility on the pass-through of the exchange rates to domestic prices in Türkiye, utilizing data from January 2006 to June 2025. Following the assessment of the stationarity levels of the variables, cointegration is identified in the model, and long-term coefficients are computed. The outputs demonstrate that a rise in exchange rates in Türkiye results in an escalation in domestic prices, and the credibility gap of the CBRT exacerbates this effect. Put differently, the degree of the pass-through of the exchange rates to domestic prices is amplified by a decline in the CBRT's credibility. The research discusses these findings.*

Keywords: Central Bank of the Republic of Türkiye; credibility; credibility gap; exchange rate pass-through; inflation.

JEL Classification: C22, E58, F41.

1. Introduction

The significance of the pass-through of exchange rates (ERs) to domestic prices, also referred to as ER pass-through (ERPT), has increased for open economies in the wake of the collapse of the Bretton Woods system, which was founded on a fixed ER regime (Faryna, 2016; Balcilar et al., 2021). Hence, there have been numerous studies conducted on this subject. The extent to which a change in ERs is reflected in domestic prices through the import price index is referred to as ERPT (Ghosh, 2013; Dedeoglu and Kaya, 2014; Anh et al., 2021; Phuc and Duc, 2021). ERPT is crucial for maintaining macroeconomic stability and informing monetary policy decisions (Kurtovic et al., 2021; Nagengast et al., 2021). In this setting, central banks (CBs) require data on the extent of ERPT to effectively address ER shocks that can induce inflationary consequences (Dedeoglu and Kaya, 2014). Certain research has indicated a decrease in ERPT for different countries (Choudhri and Hakura, 2006; Ihrig et al., 2006; Zorzi et al., 2007; Devereux and Yetman, 2010; Ha et al., 2020). Gagnon and Ihrig (2004), Mishkin and Schmidt-Hebbel (2007), and Mishkin (2008) ascribe this decline to the enhanced credibility of CBs. When a CB is firmly dedicated to maintaining low inflation, the influence of a change in ERs on domestic prices is minimal, as economic agents trust that contractionary monetary policy will be implemented by the CB when required, thereby not significantly altering their inflation expectations (IEs). Consequently, the credibility of a CB can influence ERPT (Carriere-Swallow et al., 2016).

In Türkiye, a small open economy, the inflation rate has escalated markedly, especially in recent years. This increase has significantly undermined the reputation and credibility of the Central Bank of the Republic of Türkiye (CBRT), whose principal aim is to achieve and maintain price stability. Due to the increase in inflation rates, recent research has investigated ERPT and found a significant pass-through effect from ERs to domestic prices (Tunc and Kilinc, 2018; Fendoğlu et al., 2020; Akgunduz et al., 2021; Bilgili et al., 2021, 2024; Gayaker et al., 2021; Orhan and Türel, 2022; Ilhan et al., 2023; Saygılı, 2023; Bulut et al., 2025). Among these studies, only Gayaker et al. (2021) concentrated on ERPT by incorporating the credibility of the CBRT into their model. However, this study analyzed the direct impact of credibility on the inflation rate rather than its influence on the ERPT mechanism.

This study will investigate the impact of the CBRT's credibility on ERPT by utilizing the "moderator variable" approach that is commonly employed in statistical analyses. When the interaction between two variables is also influenced by a third variable, a moderator variable is included in the empirical model (Katircioğlu and Taşpınar, 2017). Put differently, a moderator variable influences the strength of the association between the dependent variable and the regressor in a model. In this perspective, the diminishing credibility of the CBRT, especially in recent times, can influence the extent of ERPT in Türkiye. Within this scope, it is hypothesized in the study that the credibility of the CBRT can influence the extent to which the changes in ERs are transferred to domestic prices in Türkiye.

The paper utilizes monthly data over the period January 2006-July 2025. The data set commences in 2006, since the CBRT has implemented the inflation targeting (IT) regime/strategy since 2006. This study's primary strength lies in being the inaugural analysis of the influence of the CBRT's credibility on the ERPT process in Türkiye. This aspect of the investigation could make a valuable contribution to the existing body of literature.

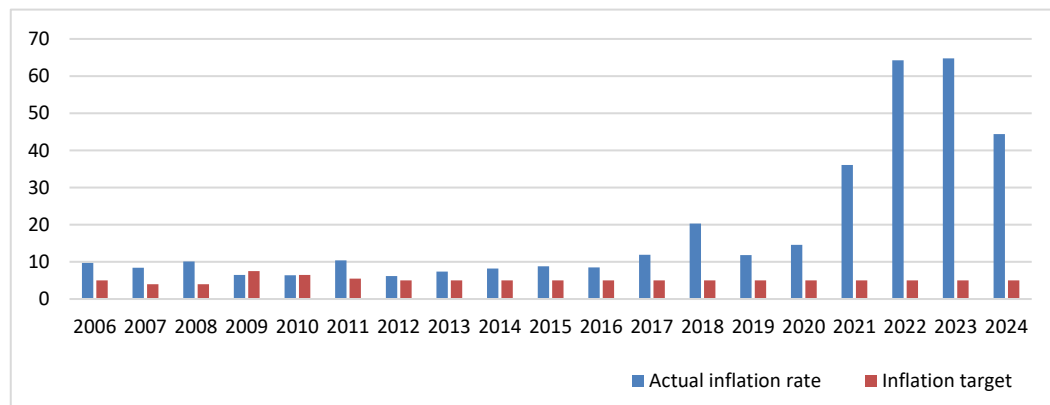
The rest of the paper is as follows: Section 2 illustrates the erosion in the credibility of the CBRT. Model and data are given in Section 3. The methodological approaches, along with the outputs of the empirical analyses, are presented in Section 4. Section 5 concludes the paper.

2. Credibility Erosion of the CBRT

Currently, there exists a near-universal agreement on implementing monetary policy within the IT strategy (Bulut, 2023). A CB that adopts the IT framework establishes a specific numerical inflation target and endeavors to achieve this target by utilizing all pertinent economic information (Svensson, 2000). The starting point for the IT strategy is the CB's ability to harmonize the public's IEs with its own inflation target, as IEs significantly influence wage negotiations and price adjustments (Carrasco and Ferreiro, 2013; Soybilgen and Yazgan, 2017). In other words, as IEs are the primary driver for inflation due to short-term price rigidities (Svensson, 1997; Bofinger et al., 2001), a CB employs monetary policy instruments to steer IEs towards the inflation target. If the CB is reputable and reliable, the inflation target will be the principal factor influencing IEs (Bulut, 2020). Hence, IEs will be minimally affected by a shock in energy prices, aggregate demand, or ERs (Bernanke, 2007; Mishkin, 2007; Gerlach et al., 2011). Within this frame, the extent to which IEs align with the inflation target is an essential measure of the CB's credibility (Çiçek and Akar, 2014). Put differently, IEs are well-anchored when they are close to the inflation target (Gürkaynak et al., 2010). In the monetary economics literature, the notion of credibility gap (CG) has been proposed to assess the credibility of a CB. Accordingly, the CG can be defined as the difference between year-end annual IEs and the year-end annual inflation target (Blinder, 2000; Lyziak, 2007).

The CBRT has implemented monetary policy under IT since 2006. An analysis of the actual inflation rates and inflation targets from 2006 to 2024 reveals that the CBRT consistently failed to meet its inflation targets. Figure 1 exhibits inflation rates and inflation targets in Türkiye during the period 2006-2024.

Figure 1: Actual inflation rates and inflation targets in Türkiye (%)

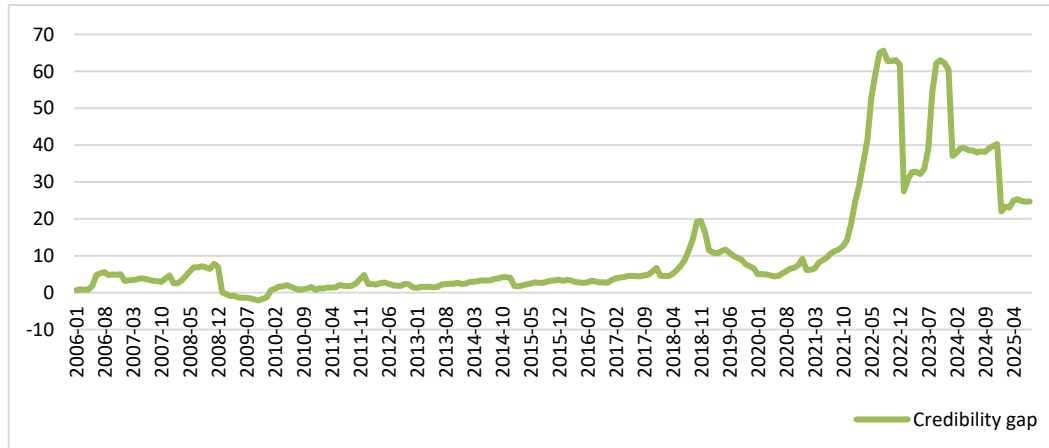


Note: Actual inflation rates and inflation targets are calculated through the consumer price index.

Source: CBRT (2025)

As is illustrated in the figure, during the period 2006-2024, only the inflation rates in 2009 and 2010 remained below the announced inflation targets. Besides, actual inflation rates have significantly deviated from the inflation targets in recent years. For instance, Türkiye's inflation target has been 5% since 2012; however, the actual inflation rates in 2022 and 2023 were 64.27% and 64.77%, respectively. Despite the fact that the inflation rate decreased to 44.38% in 2024, it remained significantly higher than the inflation target. Even though inflation is influenced by causes beyond monetary policy, resulting in CBs having limited control over it (Svensson, 1997), the CBRT, which is responsible for achieving and maintaining price stability, has been unable to achieve its inflation targets. Therefore, it is essential to evaluate the credibility of the CBRT. The CG of the CBRT for the period January 2006-August 2025 is illustrated in Figure 2. Two significant discoveries are illustrated in this figure. First, the CG of the CBRT was always positive during the observed period, indicating that IEs were higher than inflation targets. Second, there was a substantial decline in credibility, particularly in recent years, which corresponds to the notable deviation of actual inflation rates from target inflation rates.

Figure 2: *The credibility gap of the CBRT*



Notes: Credibility gap is the difference between year-end IEs calculated based on the CPI and the year-end inflation target. IEs are obtained through the survey of market participants.

Source: CBRT (2025)

Considering high inflation rates in Türkiye and the diminishing credibility of the CBRT, it is vital to investigate the possible effects of this credibility loss on the ERPT process in Türkiye. Thus, both the degree of ERPT and whether this pass-through is affected by the CBRT's credibility loss can be determined.

3. Model and Data

This research investigates the influence of the CBRT's credibility on ERPT in Türkiye. The study uses the CBRT's CG as the moderator variable and estimates an econometric model. How the CBRT's CG is used as the moderator variable can be illustrated in Figure 3.

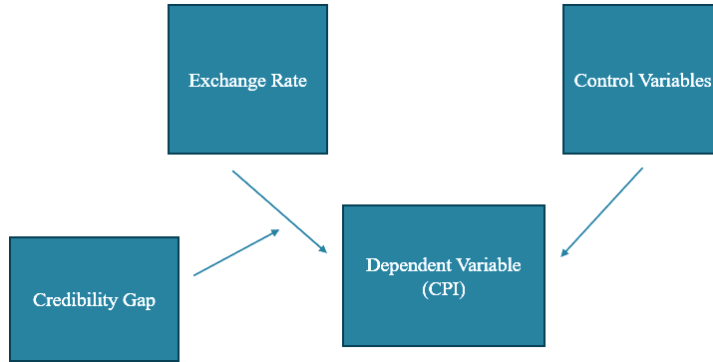
Figure 3: *The illustration of the moderator variable in the model*

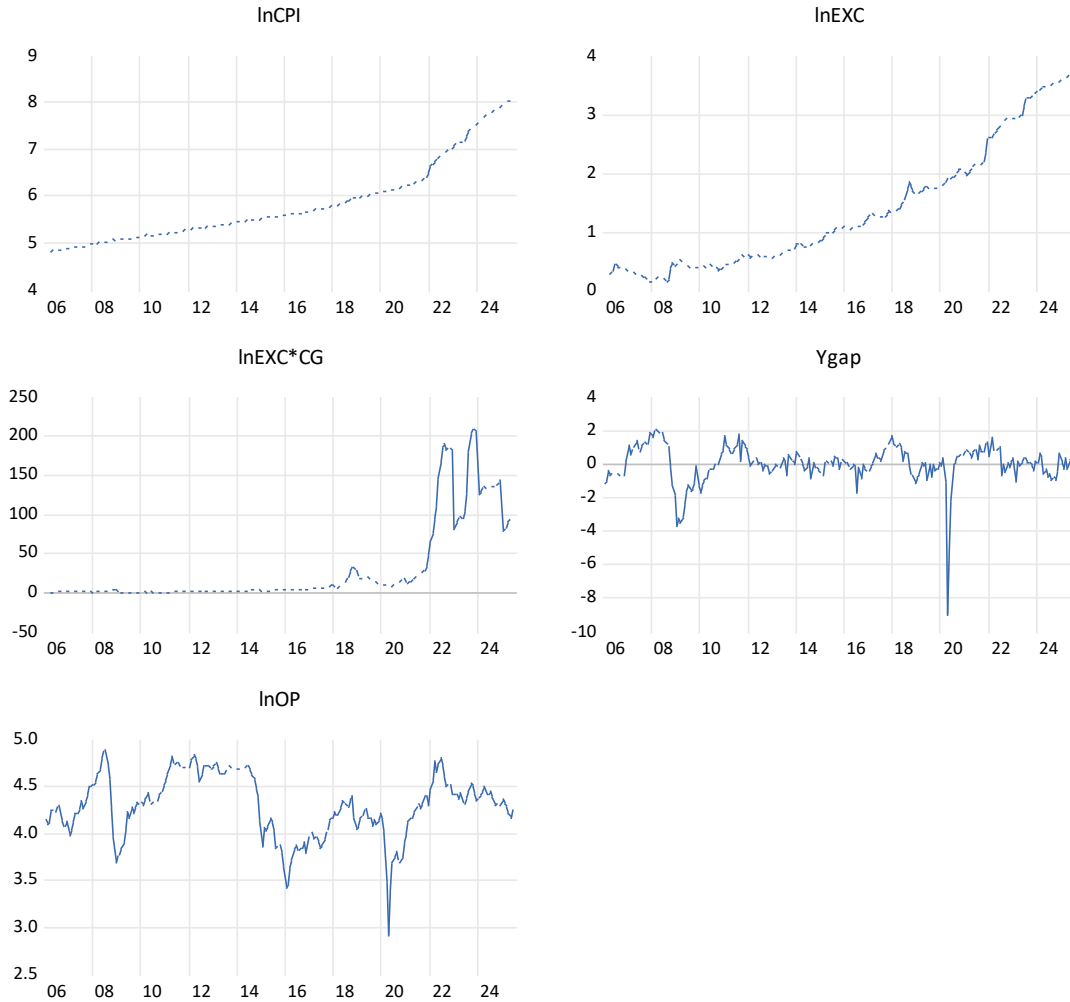
Figure 3 exhibits that the CBRT's CG serves as a moderator variable that is capable of modifying the impact of the ER on domestic prices. The study also employs output gap and oil prices as control variables in the empirical model. The extent of inflationary pressures in an economy can be deduced from the output gap, which serves as an indicator of the demand and supply components of economic activities (Jahan & Saber Mahmud, 2013).

Consequently, the output gap is integrated into the model. Given that energy sources are essential for production activities and that Türkiye predominantly imports energy to meet its demands, the model incorporates oil prices as well. Within this scope, the following model is estimated in this paper:

$$\ln \text{CPI}_t = \delta_0 + \delta_1 \ln \text{EXC}_t + \delta_2 (\ln \text{EXC}_t * \text{CG}_t) + \delta_3 Y_t^{\text{gap}} + \delta_4 \ln \text{OP}_t + \varepsilon_t \quad (1)$$

where CPI, EXC, CG, Y_t^{gap} , OP, and ε respectively denote the consumer price index (2003=100), the USD/TRY ER (TRY units per unit of USD), the CG of the CBRT, the output gap computed through the seasonally adjusted industrial production index (%), oil price (Europe Brent spot price FOB, USD per barrel), and the error term. Monthly data over the period 2006:M01-2025:M06 are utilized in the paper. Data for consumer prices and the ER are received from the CBRT (2025). The CG is calculated as the difference between the year-end IE and the year-end inflation target (see Figure 2).

The seasonally adjusted industrial production index extracted from the CBRT (2025) is filtered through the method of Hodrick and Prescott (1997) to compute the output gap. Oil price data are received from the US Energy Information Administration (2025). While the CPI, EXC, and OP series are represented in the model with their natural logarithmic values represented by $\ln$, the output gap, which is calculated as a percentage, is represented in the model with its original values. Time series dynamics of the variables under consideration are exhibited in Figure 4.

Figure 4: Time series dynamics of the variables

As is observed through Figure 3, $\ln\text{CPI}$ and $\ln\text{EXC}$, which have been increasing steadily throughout the observation period, began to increase at a faster pace in 2022. This suggests that there may be a robust correlation between the two variables. The $\ln\text{EXC}*\text{CG}$ series has experienced a significant increase in recent periods as a result of the accelerated spike in the CG. The Y^{gap} series fluctuates around zero as expected, because the level of economic activity sometimes exceeds its potential and sometimes falls below it. Similar to the movement of Y^{gap} , $\ln\text{OP}$ fluctuates around a specific average. These results suggest that the Y^{gap} and $\ln\text{OP}$ series may be stationary, whereas the other series may not be stationary.

4. Methods and Results

Before estimating coefficients, it is essential to conduct unit root tests for determining the degrees of stationarity/integration of variables in time series analysis in order to prevent

the potential spurious regression problem. The ADF unit root test of Dickey and Fuller (1981), the PP unit root test of Phillips and Perron (1988), and the ZA unit root test of Zivot and Andrews (1992) are performed in the paper. The ADF and PP unit root tests do not account for structural breaks in the series, whereas the ZA unit root test includes an endogenously determined structural break. Furthermore, all tests are based on the null hypothesis of a unit root in contrast to the alternative hypothesis of stationarity.

Table 1: ADF, PP, and ZA unit root tests

Variable	ADF		PP		ZA	
	Level	1 st dif.	Level	1 st dif.	Level	1 st dif.
lnCPI	1.572	-8.383*	2.068	-8.460*	-4.086	-8.739*
lnUSD	-0.988	-11.375*	-1.002	-9.842*	-3.068	-16.884*
lnUSD*CG	-2.052	-4.301*	-2.729	-12.760*	-6.445*	-13.204*
Y ^{gap}	-5.750*	17.425*	-5.750*	-21.273*	-6.064*	-11.512*
lnOP	-3.514**	-10.514*	-2.919	-10.538*	-4.719	-11.130*

Note: * and ** show 1% and 5% statistical significance, respectively.

The empirical outputs for the unit root tests are illustrated in Table 1. Accordingly, all tests indicate that lnCPI and lnUSD are stationary at their first difference forms, implying they are integrated of order one (I(1)). While ADF and PP tests' results reveal lnUSD*CG is stationary at its first difference form, the ZA test's finding denotes it is stationary at the level. Hence, this variable can be I(0) or I(1) based on the findings of the unit root tests. Y^{gap} is stationary at the level according to all unit root tests, meaning it is I(0). Finally, lnOP exhibits stationarity at the level with regard to the ADF unit root test, whereas the PP and ZA unit root tests' results indicate that it is stationary at its first difference value. Therefore, similar to lnUSD*CG, lnOP can be I(0) or I(1), contingent upon the results of the unit root tests.

In time series analysis, the autoregressive distributed lag (ARDL) approach to cointegration can be utilized when the dependent variable is I(1) and the regressors are either I(0) or I(1). As the results of the unit root tests strongly point out that the ARDL approximation must be utilized to search for cointegration in the established model, this study uses this technique to research whether cointegration exists in the model.

The ARDL methodology has two phases. The initial phase is to check the existence of cointegration in the model by the bounds test produced by Pesaran et al. (2001). In the second phase, if there exists cointegration in the model, the long-term parameters are computed by way of the following model (Pesaran and Shin, 1999):

$$y_t = \alpha + \sum_{i=1}^p \alpha_i y_{t-i} + \sum_{i=0}^q \beta_i x_{t-i} + u_t \quad (2)$$

The long-term coefficients are calculated via the regression model in Equation (2). Additionally, one estimates the short-run relationships via the following model depicted below:

$$\Delta y_t = \theta_0 + \theta_1 EC_{t-1} + \sum_{i=1}^p \delta_i \Delta y_{t-i} + \sum_{i=0}^q \lambda_i \Delta x_{t-i} + u_t \quad (3)$$

Eq. (3) is an error correction model. Accordingly, θ_1 indicates the extent to which short-run deviations are rectified in the long run. Consequently, it must be negative and significant to validate cointegration.

Table 2: *ARDL approach to cointegration*

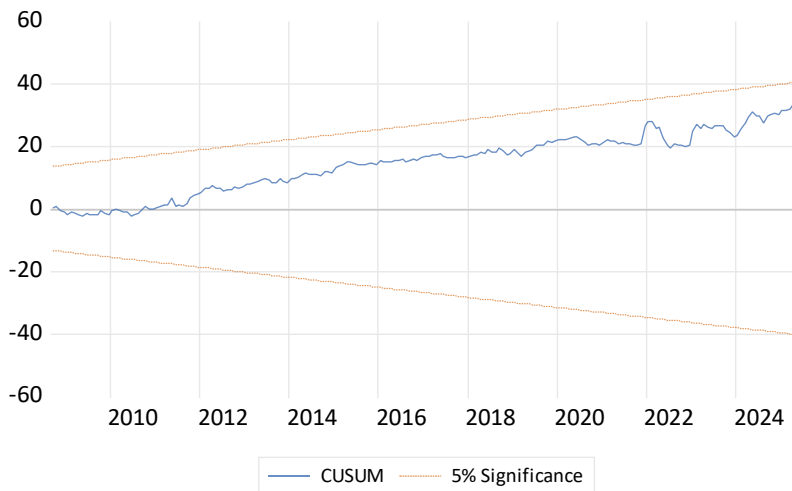
Panel A: Cointegration result			
Test stat.	4.147**		
Panel B: Coefficient estimations			
Variable	Coefficient	Standard error	t-stat.
Intercept	3.819*	0.406	9.579
lnUSD	0.816*	0.058	13.970
lnUSD*CG	0.003*	0.001	2.725
Y _{gap}	0.004	0.005	0.705
lnOP	0.226**	0.090	2.502
Panel C: Diagnostic tests			
R ² = 0.99, $\bar{R}^2$ = 0.99, F-ist. = 76111.070 (0.000), $\chi^2_{\text{BGSC}} = 0.240$ (0.886), $\chi^2_{\text{HHT}} = 28.039$ (0.214)			

Note: The error correction model's results indicate the coefficient of the one-period lagged error correction term is significant and negative. χ^2_{BGSC} and χ^2_{HHT} respectively indicate the Breusch-Godfrey serial correlation test and the Harvey heteroskedasticity test statistics. Values in parentheses indicate probability values.

* and ** show 1% and 5% statistical significance, respectively.

Table 2 demonstrates the outputs of the ARDL cointegration test. As is seen in panel A of the table, the null hypothesis of no cointegration can be rejected at 5% level, meaning that a cointegration relationship exists in the established model. This finding also implies that the long-run coefficients could be estimated. Additionally, panel B of the table demonstrates the long-run parameters for the regressors. These findings are as follows: (i) The USD/TRY ER has a positive impact on consumer prices ($\hat{\delta}_1 = 0.816$, $p < 0.01$). (ii) The variable derived from the multiplication of the USD/TRY ER and the CG of the CBRT positively affects consumer prices ($\hat{\delta}_2 = 0.003$, $p < 0.01$). (iii) Output gap has no impact on consumer prices ($\hat{\delta}_3 = 0.004$, $p > 0.10$). (iv) Consumer prices positively respond to an increase in oil prices ($\hat{\delta}_4 = 0.226$, $p < 0.05$).

The long-run parameters' stability can be assessed using the CUSUM and CUSUM-Q tests of Brown et al. (1975). It can be declared that the ARDL model produces efficient results when the CUSUM and CUSUM-Q test statistics are within the boundaries.

Figure 5: *CUSUM and CUSUM-Q tests*

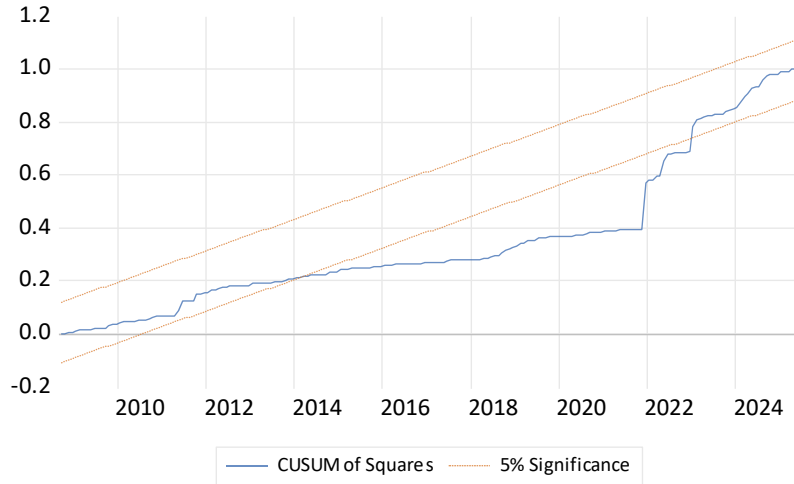


Figure 5 graphically illustrates the outcomes of the CUSUM and CUSUM-Q tests. As is seen, while the CUSUM test statistic remained within the limits throughout the observed period, the CUSUM-Q test statistic went outside the limits but later returned within the boundaries. Consequently, the results of these tests exhibit the long-term coefficient results derived from the ARDL approach are stable.

Upon estimating the coefficients using one approach, it is feasible to employ an alternative method to assess the robustness of the outputs obtained from the initial method. This research uses the dynamic ordinary least squares (DOLS) technique established by Saikkonen (1991) and Stock and Watson (1993) to evaluate the robustness of the results derived from the ARDL approach. The DOLS method is resilient to serial correlation and endogeneity issues and can be presented as follows:

$$y_t = \beta_0 + \beta_1 x_t + \sum_{i=-q}^p \gamma_i \Delta x_{t+i} + \varepsilon_t \quad (4)$$

where y stands for the dependent variable, x denotes regressors, q and p are respectively the optimal lag and lead lengths, Δ shows the first difference operator, and ε indicates the error term.

Table 3: DOLS estimator

Variable	Coefficient	Standard error	t-stat.
Intercept	4.192*	0.171	24.533
lnUSD	0.648*	0.029	21.895
lnUSD*CG	0.002*	0.001	3.880
Y^{gap}	-0.002	0.005	-0.531
lnOP	0.151*	0.037	4.072

* shows 1% statistical significance.

The results obtained through the DOLS estimator are demonstrated in Table 3. Accordingly, lnUSD, lnUSD*CG, Y^{gap} , and lnOP have the estimations of 0.648, 0.002, -0.002, BS 0.151, respectively. Besides, all coefficients except the coefficient of Y^{gap} are statistically significant. These findings indicate that consumer prices in Türkiye are positively associated with the USD/TRY ER, the variable derived from the multiplication

of the USD/TRY ER and the CG of the CBRT, and oil prices, and are not related to the output gap.

The results of the ARDL approach and the DOLS estimator imply a significant pass-through influence from the ER to domestic prices in Türkiye. Besides, they show that the CG of the CBRT amplifies the cumulative impact of the ER on domestic prices. In other words, an escalation in the CG results in heightened ERPT. The pronounced ERPT and the insignificant coefficient of the output gap suggest that Türkiye's high inflation is not fueled by demand but is instead driven by ERs and other costs. In other words, inflationary pressures fueled by demand resulting from Türkiye's rapid economic expansion do not occur. Lastly, it can be contended that the upward impact of oil prices on domestic prices is primarily due to the fact that an increase in oil prices results in an increase in the prices of petroleum products and transportation costs.

The results of the study, which suggest the presence of a robust pass-through effect in Türkiye, are corroborated by the outputs of the following studies: Tunc and Kilinc (2018), Fendoğlu et al. (2020), Akgunduz et al. (2021), Bilgili et al. (2021, 2024), Gayaker et al. (2021), Orhan and Türel (2022), İlhan et al. (2023), Saygılı (2023), and Bulut et al. (2025). Additionally, Gayaker et al. (2021) and this study demonstrate that the CG is a factor that contributes to the increase in domestic prices, even though Gayaker et al. (2021) employed a distinct methodology for modeling the impact of the CG on domestic prices.

5. Conclusion

Small open economies that are more susceptible to disruptions in international markets and cannot influence prices in those markets are particularly interested in ERPT. Recent research has revealed that the extent of this pass-through is reduced when a CB has a high level of credibility. These results are attributable to the public's confidence in the CB's ability to preserve price stability. Accordingly, the CB can manage IEs against an increase in ERs. Consequently, it can be argued that ERPT can be influenced by the CB's credibility.

In recent years, inflation rates in Türkiye have increased significantly, and the CBRT has encountered significant challenges in managing IEs. Consequently, this study investigates the extent to which the ER passes through to domestic prices in Türkiye using monthly data from 2006:M01 to 2025:M06. The CG of the CBRT, which can influence the interaction between domestic prices and the ER, is employed as a moderator variable in this context. After performing the unit root tests, which indicate that the variables have different orders of integration, the cointegration relationship between the variables is examined. Following the discovery of cointegration, long-run parameters are estimated.

The results indicate that an increase in the ER results in an increase in domestic prices and that the CG of the CBRT amplifies this effect. In other words, ERPT is exacerbated by a decrease in the CBRT's credibility. Therefore, the less confidence there is that the CBRT will meet its inflation target, the stronger the pass-through of the ER to domestic prices becomes in Türkiye. It is widely accepted that if IEs are well anchored due to a highly credible CB, the reaction of domestic prices to ER shocks will lessen (Mishkin and

Schmidt-Hebbel, 2007). Besides, in a country with high CB credibility, ERPT can be low, which makes it easier for monetary policy to control inflation (Carriere-Swallow et al., 2016). In this way, the strong commitment of a credible CB to sustaining low and stable inflation has been a crucial element in the limited transmission of the ER shocks to domestic prices over the last decades (Ha et al., 2020). However, in Türkiye, where the CBRT has a serious credibility problem and IEs are not well-anchored, the erosion of the CB's credibility intensifies the impact of an increase in ERs on domestic prices, contributing to high inflation rates. Consequently, this paper argues that the CBRT should (i) employ its monetary policy instruments in accordance with its inflation target, (ii) manage IEs much more effectively, and (iii) endeavor to align IEs with the inflation target. Besides, to mitigate the escalation of domestic prices resulting from an increase in the ER due to heightened prices of imported inputs and final goods and services, policymakers in Türkiye should implement policies that can reduce the Turkish economy's dependence on imports.

Conflict of interest

The authors declare that they have no known competing interests or personal relationships that could have appeared to influence the work reported in this paper.

Original Submission Statement

The manuscript has not been previously published in its current form or a substantially similar form and is not currently under review by another journal.

Data availability: The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Funding: This research was funded by TÜBİTAK under project number 1919B012411388.

Acknowledgements: Not applicable

List of Abbreviations

Abbreviations	Explanation
ADF	Augmented Dickey-Fuller
ARDL	Autoregressive distributed lag
CB	Central bank
CBRT	Central Bank of the Republic of Turkey
CG	Credibility gap
DOLS	Dynamic ordinary least squares
ER	Exchange rate
ERPT	Exchange rate pass-through
IE	Inflation expectation
IT	Inflation targeting
PP	Phillips-Perron
ZA	Zivot-Andrews

References

- Akgunduz, Y.E., Bastan, E.M., Demiroglu, U., and Tumen, S. (2021). Product-level estimates of exchange rate pass-through: evidence from Turkey. *The World Economy*, 44(7), 2203-2226.
- Anh, V.T., Quan, L.T.T., Phuc, N. van, Chi, H.M., and Duc, V.H. (2021). Exchange rate pass-through in ASEAN countries: an application of the SVAR model. *Emerging Markets Finance and Trade*, 57(1), 21-34.
- Balcilar, M., Roubaud, D., Usman, O., and Wohar, M.E. (2021). Testing the asymmetric effects of exchange rate pass-through in BRICS countries: does the state of the economy matter? *World Economy*, 44(1), 188-233.
- Bernanke, B.S. (2007). Inflation expectations and inflation forecasting. <<https://www.federalreserve.gov/newsevents/speech/bernanke20070710a.htm>> [Accessed September 1, 2025].
- Bilgili, F., Ünlü, F., Gençoğlu, P., and Kuşkaya, S. (2021). Modeling the exchange rate pass-through in Turkey with uncertainty and geopolitical risk: a Markov regime-switching approach. *Applied Economic Analysis*, 30(88), 52-70.
- Bilgili, F., Kuşkaya, S., Gençoğlu, P., Ünlü, F., and Magazzino, C. (2024). Exchange rate pass-through at high and low frequencies in Turkey: a continuous wavelet transform model approach. *Journal of Economic Structures*, 13(1), 21.
- Blinder, A.S. (2000). Central-bank credibility: why do we care? How do we build it? *American Economic Review*, 90(5), 1421-1431.
- Bofinger, P., Reischle, J., and Schachter, A. (2001). *Monetary Policy: goals, institutions, strategies, and instruments*. New York: Oxford University Press.
- Brown, R.L., Durbin, J., and Evans, J.M. (1975). Techniques for testing the constancy of regression relationships over time. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 37(2), 149-163.
- Bulut, U. (2020). The credibility problem of the Central Bank of the Republic of Turkey: An empirical investigation (in Turkish). *Bulletin of Economic Theory and Analysis*, 5(2), 1-15.
- Bulut, U. (2023). Measuring the impacts of monetary policy in Turkey: an extended structural vector autoregressive model with structural breaks. *International Journal of Economic Policy Studies*, 17(1), 117-132.
- Bulut, U., Togay, S., and Ertugrul, H.M. (2025). Price rigidities, central bank credibility, and exchange rate pass-through: the case of Türkiye. *Ekonomik Yaklasim*, 36(136), 267-291.
- Carrasco, C.A., and Ferreira, J. (2013). Inflation targeting and inflation expectations in Mexico. *Applied Economics*, 45(23), 3295-3304.
- Carriere-Swallow, M.Y., Gruss, B., Magud, M.N.E., and Valencia, M.F. (2016). Monetary policy credibility and exchange rate pass-through. *IMF Working Paper*, 240.
- Central Bank of the Republic of Türkiye (2025). Electronic data delivery system. <<https://evds2.tcmb.gov.tr/>> [Accessed September 15, 2025].
- Choudhri, E.U., and Hakura, D.S. (2006). Exchange rate pass-through to domestic prices: does the inflationary environment matter? *Journal of International Money and Finance*, 25(4), 614-639.
- Çiçek, S., and Akar, C. (2014). Do inflation expectations converge toward inflation target or actual inflation? evidence from expectation gap persistence. *Central Bank Review*, 14, 15-21.

- Dedeoglu, D., and Kaya, H. (2014). The evolution of exchange rate pass-through in Turkey: does inflation targeting matter? *Afro Eurasian Studies*, 3(1), 26–33.
- Devereux, M.B., and Yetman, J. (2010). Price adjustment and exchange rate pass-through. *Journal of International Money and Finance*, 29(1), 181–200.
- Dickey, D., and Fuller, W. (1981). Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica*, 49(4), 1057–1072.
- Energy Information Administration (2025). Petroleum & other liquids. <<https://www.eia.gov/dnav/pet/hist/RBRTED.htm>> [Accessed September 15, 2025].
- Faryna, O. (2016). Nonlinear exchange rate pass-through to domestic prices in Ukraine. *Visnyk of the National Bank of Ukraine*, 236, 30–42.
- Fendoğlu, S., Çolak, M.S., and Hacıhasanoğlu, Y S. (2020). Foreign-currency debt and the exchange rate pass-through. *Applied Economics Letters*, 27(8), 657–666.
- Gagnon, J.E., and Ihrig, J. (2004). Monetary policy and exchange rate pass-through. *International Journal of Finance and Economics*, 9(4), 315–338.
- Gayaker, S., Ağaslan, E., Alkan, B., and Çiçek, S. (2021). The Deterioration in credibility, destabilization of exchange rate and the rise in exchange rate pass-through in Turkey. *International Review of Economics and Finance*, 76(June 2020), 571–587.
- Gerlach, P., Hordahl, P., and Moessner, R. (2011). Inflation expectations and the Great Recession. *BIS Quarterly Review*, 39-51.
- Ghosh, A. (2013). Exchange rate pass through, macro fundamentals and regime choice in Latin America. *Journal of Macroeconomics*, 35(1), 163–171.
- Gürkaynak, R.S., Levin, A., and Swanson, E. (2010). Does inflation targeting anchor long-run inflation expectations? Evidence from the US, UK, and Sweden. *Journal of the European Economic Association*, 8(6), 1208-1242.
- Ha, J., Stocker, M.M., and Yilmazkuday, H. (2020). Inflation and exchange rate pass-through. *Journal of International Money and Finance*, 105, 102187.
- Hodrick, R.J., and Prescott, E.C. (1997). Postwar US business cycles: an empirical investigation. *Journal of Money, Credit, and Banking*, 29(1), 1-16.
- Ihrig, J.E., Marazzi, M., and Rothenberg, A.D. (2006). Exchange rate pass-through in G-7 countries, Federal Reserve System, *International Finance Discussion Paper* 851.
- Ilhan, A., Akdeniz, C., and Özdemir, M. (2023). Time-varying exchange rate pass-through to domestic prices: Evidence from Turkey. *Romanian Journal of Economic Forecasting*, 26(3), 162-182.
- Katircioğlu, S.T., and Taşpınar, N. (2017). Testing the moderating role of financial development in an environmental Kuznets curve: empirical evidence from Turkey. *Renewable and Sustainable Energy Reviews*, 68, 572-586.
- Kwiatkowski, D., Phillips, P.C., Schmidt, P., and Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? *Journal of Econometrics*, 54(1-3), 159-178.
- Lyziak, T., Mackiewicz, J., and Stanislawska, E. (2007). Central bank transparency and credibility: the case of Poland, 1998-2004. *European Journal of Political Economy*, 23(1), 67-87.
- Mishkin, F.S. (2007). Inflation dynamics. *International Finance*, 10 (3), 317-334.
- Mishkin, F.S., and Schmidt-Hebbel, K. (2007). Does inflation targeting make a difference? *NBER Working Paper*, 12876.

- Mishkin, F.S. (2008). Exchange rate pass-through and monetary policy, *NBER Working Paper*, 13889.
- Orhan, A., and Türel, M. (2022). Asymmetries in exchange rate pass-through in Turkey: A threshold VAR analysis. *Prague Economic Papers*, 31(3-4).
- Pesaran, M.H., and Shin, Y. (1999). "An Autoregressive Distributed Lag Modelling Approach to Cointegration Analysis", In *Econometrics and econometric theory in the 20th century: the Ragnar Frisch centennial symposium*, ed. S. Strom, 371-413. Cambridge: Cambridge University Press.
- Pesaran, M.H., Shin, Y., and Smith, R.J. (2001). bounds testing approaches to the analysis of level relationships. *Journal of Applied Econometrics*, 16(3), 289–326.
- Phillips, P.C.B., and Perron, P. (1988). Testing for a unit root in time series regression. *Biometrika*, 75(2), 335–346.
- Phuc, N. van, and Duc, V.H. (2021). Macroeconomics determinants of exchange rate pass-through: new evidence from the Asia-Pacific region. *Emerging Markets Finance and Trade*, 57(1), 5–20.
- Saikkonen, P. (1991). Asymptotically efficient estimation of cointegration regressions. *Econometric Theory*, 7(1), 1-21.
- Saygılı, H. (2023). Invoicing currency, exchange rate pass-through, and value-added trade: The case of Turkey. *International Journal of Finance & Economics*, 28(4), 4401-4419.
- Soybilgen, B., and Yazgan, E. (2017). An evaluation of inflation expectations in Turkey. *Central Bank Review*, 17(1), 31–38.
- Stock, J. H., and Watson, M. W. (1993). A simple estimator of cointegrating vectors in higher order integrated systems. *Econometrica*, 61, 783-820.
- Svensson, L.E. (1997). Inflation forecast targeting: Implementing and monitoring Inflation targets. *European Economic Review*, 41(6), 1111-1146.
- Svensson, L.E. (2000). How should monetary policy be conducted in an era of price stability?". *NBER Working Paper*, 7516.
- Tunc, C., and Kilinc, M. (2018). Exchange rate pass-through in a small open economy: a structural VAR approach. *Bulletin of Economic Research*, 70(4), 410–422.
- Zorzi, M.C., Hahn, E., and Sanchez, M. (2007). Exchange rate pass-through in emerging markets. *European Central Bank Working Paper*, 739.

The aggregate banking-sector footprint of financial education mandates: cohort exposure, community bank margins, and an arithmetic of detection

Dan-Mihai DIMA

Bucharest University of Economic Studies, Romania
dan.dima@fin.ase.ro

Abstract. *We measure the community-bank footprint of state financial education mandates using staggered adoption across seven US states (2005–2019). The treatment is a continuous Cohort Exposure Index (CEI) capturing the mandate-exposed share of the working-age population. In a panel of 81,549 community bank-year observations, CEI is positively associated with net interest margin; the gradient is concentrated in banks with higher consumer loan share and is accompanied by elevated consumer-loan charge-offs. An arithmetic decomposition bounds the channel's expected per-bank magnitude; our estimates fall inside this band, consistent with the bank-balance-sheet channel operating at the scale the underlying cohort mechanism allows.*

Keywords: financial literacy, financial education mandates, net interest margin, community banks, staggered difference-in-differences.

JEL Classification: G21, G28, I22, H75.

1. Introduction

Since 2000, more than half of US states have enacted laws requiring high school students to complete a standalone personal finance course as a condition of graduation. Household-level evidence shows that students exposed to these mandates display higher credit scores, lower delinquency, and a shift toward formal lending products by their early twenties (Brown et al., 2016; Stoddard and Urban, 2020; Burke et al., 2023). Whether these household-level shifts aggregate into the balance sheets of the intermediaries that serve the same populations is an open question. This paper asks it for community banks: the institutions most plausibly exposed to the cohort mechanism through their consumer-credit franchises.

The empirical strategy exploits staggered mandate adoption across seven states between 2008 and 2020. The treatment variable is a continuous Cohort Exposure Index (CEI) constructed at the state-year level: the share of the working-age population whose secondary schooling fell on or after the state's mandate adoption year, computed under a uniform age-distribution approximation across the working-age population. The CEI rises gradually as successive mandate-exposed cohorts age into the adult population, providing a continuous dose measure rather than a binary on-off treatment indicator. The dependent variable is the net interest margin (NIM) of community banks operating in the state, drawn from FDIC Call Reports. The panel covers 81,549 bank-year observations on 7,841 banks.

In the baseline two-way fixed-effects specification, CEI is positively associated with NIM with a marginally significant coefficient that depends on the inclusion of the Global Financial Crisis (GFC) window. The more informative result comes from a heterogeneity specification that interacts CEI with the bank's consumer loan share (CLS): the interaction is positive and economically large ($\hat{\beta}_3 = 11.46$), and a Fisher permutation test under geographic placebo reassignment places it well above the placebo null distribution (one-sided empirical $p = 0.002$). State-clustered asymptotic inference is imprecise given the small treated-cluster count, which is why we treat the permutation p -value as the primary inferential criterion. A supplementary charge-off regression delivers a significantly positive $\text{CEI} \times \text{CLS}$ interaction, indicating that consumer-credit concentrated banks in high-exposure states earn higher NIMs while also absorbing correspondingly higher credit losses.

Two interpretive issues structure the rest of the paper. The first is whether these estimates are large or small relative to the mechanism that generates them. We address this with an arithmetic decomposition of the expected per-bank NIM footprint into three multiplicative factors: the treated-population share, the exposed-portfolio share, and the household-level yield differential. The decomposition delivers an implied band rather than a point ceiling; our estimates fall inside the band, which we read as the bank-balance-sheet channel operating at a scale consistent with, not above or below, the underlying cohort mechanism. The second issue is what the design can and cannot deliver. With seven treated states, two of which adopted during the GFC, the design has real power for the within-state

heterogeneity gradient but limited power for the aggregate dose-response coefficient. Sharper identification of the same mechanism requires data closer to the unit of behavior, such as individual credit-bureau records or product-level loan portfolios; the CEI construction transfers directly to those settings.

The paper makes two contributions. The Cohort Exposure Index provides a continuous dose measure for any cohort-based education or behavioral policy whose effects diffuse through demographic replacement; it avoids the errors-in-variables problem that affects discrete treatment-timing indicators when the underlying mechanism operates on a fraction of the population. The arithmetic decomposition provides an ex-ante calibration tool for evaluating cohort-based policies at aggregate intermediary scale, and an interpretive bridge between household-level and balance-sheet-level evidence on the same mechanism.

Section 2 reviews the literature. Section 3 describes the identification strategy and the CEI construction. Section 4 describes the data and summary statistics. Section 5 presents the empirical results. Section 6 develops the arithmetic-of-detection decomposition. Section 7 concludes.

2. Literature Review

2.1. From Individual Financial Behavior to Aggregate Market Outcomes

Individual financial literacy shapes household-level financial outcomes along several margins. Households with higher financial literacy hold more diversified portfolios, accumulate greater retirement wealth, incur lower debt costs, and are less likely to carry high-interest revolving balances (Lusardi and Mitchell, 2014). Whether the aggregation of individual financial competence across households produces a corresponding effect on the efficiency of financial intermediation is less well documented.

The micro-literacy literature analyses behavioral effects at the household or firm level; the macro-financial-development literature treats banking efficiency as a function of institutional quality, competition, regulatory design, and macroeconomic conditions. This paper exploits within-country, subnational variation in exposure to mandated financial education and links that variation to the net interest margins of community banks operating in the same local markets.

2.2. Financial Literacy, Financial Inclusion, and Macro-Financial Development

Grohmann et al. (2018) provide the first large-scale cross-country evidence that household financial literacy improves financial inclusion, using the 2014 S&P Global Financial Literacy Survey across 148 countries combined with World Bank Global Financial Development Database indicators. Hussain et al. (2021) extend this framework to a 98-country panel and document that the literacy–inclusion channel is strongly moderated by institutional quality. Basha et al. (2023) shift the unit of analysis to firms and show that small-firm leverage responds to the interaction of financial literacy and financial

development. Sethi and Acharya (2018) document that financial inclusion Granger-causes economic growth in a 31-country panel. Hussain et al. (2024) find substantial asymmetry in the inclusion–growth relationship across developed and developing Asian economies.

A contrary finding comes from Barik and Pradhan (2020), who study the BRICS economies and document that rapid financial inclusion expansion driven by credit-standard relaxation is associated with deteriorating financial stability. The mechanism we document in Section 5.5 is closer in spirit to Barik than to a pure asset-quality narrative: financial-education-driven credit expansion in the present setting is accompanied by, not in place of, elevated charge-offs. Nogueira et al. (2025) reverse the standard causal arrow, finding that per-capita income, financial system development, and educational attainment are themselves primary drivers of aggregate financial literacy across 20 OECD countries, a concern for endogenous mandate adoption that our staggered identification design addresses.

2.3. Micro-Foundations for an Intermediation Efficiency Channel

Four theoretical channels link the CEI to bank NIMs. (i) A deposit-side channel: financially illiterate depositors exhibit high search costs and inertia, allowing banks to earn rents on liabilities; rising literacy makes deposit supply more elastic and compresses NIMs (Erel et al., 2023). (ii) A loan-side quality channel: financially literate borrowers reduce adverse selection and moral hazard, lowering expected credit losses and the default risk premium. (iii) A competition channel: illiterate borrowers underestimate costs and refinance less frequently (Stango and Zinman, 2009; Woodward and Hall, 2012), allowing banks to extract pricing premia that rising literacy erodes. (iv) A credit-demand channel: literate adults who understand credit products may be more willing to borrow, acting as a demand stimulus rather than only a source of risk reduction. Under (iv), community banks in high-literacy markets face an expanded pool of informed retail borrowers, may extend their consumer portfolios, and may capture the yield premium of installment credit—generating NIM expansion alongside, rather than instead of, elevated expected credit losses. The empirical evidence in Section 5.5 is consistent with (iv).

2.4. Where the Present Paper Sits

Two literatures bracket this paper. The household-side work cited above (Brown et al., 2016; Stoddard and Urban, 2020; Burke et al., 2023; Mangrum, 2022) establishes that mandate-exposed cohorts exhibit measurable changes in credit behavior, identified from individual credit-bureau records. The macro-financial-development literature treats banking efficiency as a function of aggregate institutional and macroeconomic conditions. We sit between them. The empirical question is whether the household-level shifts already documented aggregate into a detectable response on the balance sheets of the intermediaries that serve the same populations, and at what magnitude. The CEI provides the continuous dose measure that the dose-response question requires; the heterogeneity specification isolates the segment of the banking sector where the channel can plausibly operate.

3. Methodology and Identification Strategy

3.1. Research Design Overview

The empirical strategy exploits the staggered, quasi-random timing with which US states adopted standalone personal finance graduation requirements between 2008 and 2020. The primary outcome is the net interest margin (NIM) of FDIC-insured community banks operating in state s in year t . The identifying assumption is that, conditional on pre-mandate covariates and bank fixed effects, the timing of mandate adoption is uncorrelated with pre-existing trends in bank NIMs—an assumption tested formally using pre-period event-study estimates.

3.2. The Treatment Variable

The treatment indicator is $D_{s,t} \in \{0,1\}$, equal to one for all years t on or after the year in which state s enacted a standalone personal finance graduation requirement. Adoption years are coded using the NGPF (Next Gen Personal Finance) Live Dashboard as the primary source, cross-referenced with the Council for Economic Education's biennial Survey of the States. Treatment cohorts in the 2005–2019 window: Utah (2008), Virginia (2009), Tennessee and Alabama (2013–2014), Missouri (≈ 2017), North Carolina (2019), and Iowa (≈ 2020). We define G_s as the calendar year of first treatment, with $G_s = \infty$ for the approximately 44 never-treated states. Missouri and Iowa are coded with approximate adoption years; alternative codings within the cited primary-source range leave the main estimate essentially unchanged.

3.3. The Cohort Exposure Index

The binary indicator $D_{s,t}$ does not capture the dosage of financial education received by the adult population at time t . A mandate adopted in year g affects only students enrolled in high school during and after that year; the mandate-exposed share of the adult population increases gradually as successive cohorts age into adulthood. We construct the Cohort Exposure Index (CEI), a continuous dose variable. An individual of age a residing in state s in year t is classified as exposed if their high school years overlapped with the period in which the mandate was in effect:

$$\mathbf{1}[\text{exposed}(a, s, t)] = \mathbf{1}[t - a + 17 \geq G_s] \times \mathbf{1}[G_s < \infty] \quad (1)$$

The CEI for state s in year t is the exposure share of the 18–65 working-age population:

$$\text{CEI}_{s,t} = \sum_{a=18}^{65} \omega_{a,s,t} \times \mathbf{1}[\text{exposed}(a, s, t)] \quad (2)$$

where $\omega_{a,s,t}$ is the population weight assigned to age a in state s in year t . In principle, this weight could be sourced from Census Bureau Population Estimates Program (PEPS) and American Community Survey single-year-of-age estimates by state and year (vintages 2000–2009 and 2010–2019) to let the exposure share reflect each state's actual age distribution; we leave that refinement to future work with finer administrative age data. The specification estimated throughout this paper instead assumes a uniform age

distribution across the working-age bracket, under which Equation (2) reduces to the closed-form expression:

$$CEI_{s,t} = \frac{\max(0, \min(48, t - G_s))}{48} \quad (3)$$

which is the specification we estimate throughout the paper; Section 5.4.3 discusses this construction choice and a demographic-weighted extension. By construction, $CEI_{s,t} \in [0,1]$, equals zero for all never-treated states, and increases monotonically in treated states at approximately $1/48 \approx 2.08$ percentage points per year as each successive mandate-exposed high school cohort ages into the working-age population. The maximum observed value is 0.229 (Utah in 2019, eleven years post-adoption).

3.4. The Community Bank Restriction

We restrict the analysis to FDIC-insured commercial banks and savings institutions with year-end total assets below \$1 billion (BKCLASS codes N, NM, SM, SA, and SB). This threshold is the standard definition used by the FDIC for the community banking sector (Federal Deposit Insurance Corporation, 2020). Large multi-state banking organizations are excluded because their NIMs are determined at the national portfolio level by centralized asset-liability management, rendering any single state's financial literacy shock trivial relative to the national book.

3.5. Estimation Framework

The identification strategy rests on a continuous-treatment Two-Way Fixed Effects (TWFE) design. The treatment variable, the Cohort Exposure Index, is a continuous dosage measure that accumulates as mandate-exposed cohorts age into the working-age population. The estimand of interest is therefore the marginal NIM response to a one-unit increase in CEI rather than a binary average treatment effect.

Recent methodological work raises concerns about TWFE in staggered settings under heterogeneous treatment effects (Callaway and Sant'Anna, 2021; Sun and Abraham, 2021). Two features of our setting mitigate these concerns. First, the continuous CEI specification does not require cohort-specific homogeneity in the same way as a binary event study: the dose-response slope is identified from within-state variation in cumulative exposure. Second, with only seven treated states, we do not rely on asymptotic clustered inference; instead we apply a non-parametric Fisher permutation test that re-assigns observed mandate adoption years to randomly drawn control states and re-estimates the full regression 1,000 times, generating an exact finite-sample null distribution robust to both staggered timing and the small treated-cluster count. We treat the permutation p -value as the primary inferential criterion alongside the continuous TWFE estimates.

3.6. Main Specification

The primary TWFE regression is:

$$NIM_{b,s,t} = \alpha_b + \delta_t + \beta CEI_{s,t} + \gamma' X_{s,t} + \varepsilon_{b,s,t} \quad (4)$$

where α_b are bank fixed effects, δ_t are year fixed effects, and $\mathbf{X}_{s,t}$ includes the state unemployment rate and real GDP growth. Standard errors are clustered at the state level. The heterogeneity specification augments Equation (4) with an interaction term:

$$\text{NIM}_{b,s,t} = \alpha_b + \delta_t + \beta_1 \text{CEI}_{s,t} + \beta_2 \text{CLS}_{b,t} + \beta_3 \text{CEI}_{s,t} \times \text{CLS}_{b,t} + \gamma' \mathbf{X}_{s,t} + \varepsilon_{b,s,t} \quad (5)$$

where $\text{CLS}_{b,t}$ is the bank's consumer loan share (LNCON/LNLSNET, consumer installment and revolving credit excluding residential mortgages), and $\beta_1, \beta_2, \beta_3$ are the coefficients on CEI, CLS, and their interaction. The marginal effect of mandate exposure on NIM is $\partial \text{NIM} / \partial \text{CEI} = \hat{\beta}_1 + \hat{\beta}_3 \times \text{CLS}_{b,t}$.

The choice of consumer loan share as the primary heterogeneity variable follows directly from the life-cycle stage of the mandate-exposed cohort. Financial literacy mandates target students at ages 15–17 (high school); those graduates enter the labor market at approximately 18–22 and reach peak consumer credit-market entry by their mid-twenties. The first credit products for this cohort are overwhelmingly auto loans, credit cards, and personal installment loans—not residential mortgages. The median age of first-time homebuyers in the United States is approximately 35 years, well outside the 18–25 window that defines the mandate-exposed cohort in our sample horizon. Consequently, the relevant portfolio-mix dimension for identifying the mandate transmission channel is the share of *consumer* credit in a bank's loan book, not the combined retail share that includes residential mortgages driven by older, unaffected cohorts. As a robustness check we also estimate the model with the broader retail loan share ($\text{RLS} = (\text{LNRERES} + \text{LNCON}) / \text{LNLSNET}$), which adds 1–4 family mortgages to the consumer installment component; the interaction estimate is +6.882 (SE = 5.152), directionally consistent with the CLS specification and smaller in magnitude, as expected if mortgages dilute the consumer-demand signal.

3.7. Threats to Identification

Endogenous mandate adoption. Early adopters are systematically different from non-adopters along observable dimensions: states that enacted financial literacy requirements first tend to have above-median household incomes, denser retail bank branch networks, higher pre-mandate credit penetration rates (outstanding consumer and mortgage debt per capita), and stronger baseline educational attainment. Each of these characteristics is plausibly correlated with both the propensity to adopt early *and* the level of community bank NIM, creating a selection threat. The Callaway and Sant'Anna (2021) estimator addresses this through conditional parallel trends – allowing trend comparisons only within cells matched on pre-treatment covariates – but implementing it rigorously here would require conditioning on state-level pre-mandate household income, consumer credit penetration, and branch density, which we do not do.

Instead, we rely on two sources of reassurance. First, the event-study pre-trend test in Figure 1 yields $F = 7.95$ ($p = 0.159$) on the five pre-adoption periods: if early-adopting states were on systematically different NIM growth trajectories before their mandates, we

would expect negative and significant pre-event coefficients, but we find none. Second, CEI varies continuously within treated states over time, so identification exploits the gradual demographic accumulation of mandate-exposed cohorts rather than the single adoption event – a within-state temporal source of variation that is largely orthogonal to the static cross-sectional differences that drive selection.

We flag this as a limitation and acknowledge that formal covariate conditioning remains an open robustness exercise. **COVID-19 macro confound.** Restricting the sample to $t \leq 2019$ eliminates the confounding effects of near-zero policy rates and emergency asset purchase programs. **Multi-state bank contamination.** The community bank restriction ensures each institution's NIM reflects pricing conditions in the treatment state's local market.

State-specific characteristics of early adopters. Two of the seven treated states, Utah (2008) and Virginia (2009), warrant specific discussion given features that may affect the generalizability of the estimates. Virginia is geographically proximate to Washington, D.C. and receives a disproportionately high inflow of residents from other states due to federal government employment and defense contracting. This high mobility rate means that the CEI may overstate the fraction of Virginia bank customers actually educated under the Virginia mandate, as a substantial share of working-age adults grew up elsewhere. To the extent that this dilutes the measured treatment intensity, it biases our estimates toward zero and implies that the true mandate effect may be larger than reported. A Virginia-excluded leave-one-out specification yields a larger positive interaction estimate (+17.21, SE = 5.43, $p = 0.002$), consistent with attenuation-to-zero rather than a spurious Virginia-specific effect.

Utah's population is predominantly affiliated with the Church of Jesus Christ of Latter-day Saints, whose financial norms emphasize debt avoidance and self-reliance. Relative to a national benchmark, this may imply a lower baseline demand for consumer credit in Utah, which could attenuate the mandate-to-credit-demand mechanism. However, the relevant margin is whether *within-Utah* variation in mandate exposure changes consumer borrowing behavior—not whether Utah residents are above- or below-average borrowers in an absolute sense. The bank-level fixed effects absorb Utah's persistent low-credit-demand level; the CEI identifies effects from the *change* in the educated share of the Utah population over time. That said, we acknowledge that the estimated effect for Utah may differ from what would be expected in a more nationally representative state. Dropping Utah from the leave-one-out specification reverses the interaction estimate to -9.15 (SE = 2.54, $p < 0.001$), confirming Utah's outsized leverage in the full-sample estimate; the Fisher permutation test is not subject to this leverage concern.

Inference with small treated-cluster count. With only seven treated states, asymptotic clustered standard errors are unreliable (Cameron et al., 2008; MacKinnon and Webb, 2017). Our primary inferential criterion is therefore the Fisher permutation test (empirical one-sided $p = 0.002$), which does not rely on asymptotics. The permutation design

randomizes mandate adoption across all 51 state jurisdictions, so the null distribution reflects the full cross-state heterogeneity in community bank characteristics—not merely the characteristics of VA and Utah. A result that survives this benchmark is unlikely to be an artifact of any single state’s idiosyncrasies.

4. Data and Summary Statistics

4.1. Data Sources

The empirical analysis combines three data sources.

Bank financial data. The FDIC BankFind Suite financials endpoint provides December 31 annual call-report observations for 2005–2019. We apply BKCLASS restrictions (codes N, NM, SM, SA, and SB) and the \$1 billion total-asset ceiling. The dependent variable is NIMY (annualized net interest margin, percent). The primary heterogeneity variable is the **consumer loan share**: $CLS_{b,t} = LNCON/LNLSNET$, the ratio of consumer installment and revolving loans (auto, credit card, and personal installment credit) to total net loans and leases, *excluding residential mortgages*. This definition directly matches the credit products demanded by mandate-exposed cohorts (ages 18–25): the average age of a first-time homebuyer is approximately 35 years, so mortgage demand in our sample period is driven primarily by older, non-exposed cohorts. Consumer credit, by contrast, is the primary financial product for young adults entering the labor market with enhanced financial knowledge.

Among the seven treated states, community banks hold a median consumer loan share of approximately 4.2 percent of net loans (mean 6.7 percent, interquartile range 1.7–8.4 percent), reflecting meaningful cross-sectional heterogeneity in business-model orientation toward consumer credit. As a robustness check that tests sensitivity to the mortgage-exclusion decision, we also report results using the broader retail loan share $RLS_{b,t} = (LNRERES + LNCON)/LNLSNET$, which adds 1–4 family residential mortgages to the consumer component.

Community bank market presence. To gauge the relevance of the community bank channel, we compute the share of total insured deposit volume held by community banks (total assets < \$1B) within each treated state using the FDIC Summary of Deposits. Across the seven treated states, community banks account for between 12 and 35 percent of total insured deposits, with a cross-state median of approximately 22 percent. While smaller than the aggregate market shares of large banks, this presence is economically meaningful: community banks are disproportionately concentrated in retail consumer and mortgage lending relative to their deposit share, so their responsiveness to local borrower composition shocks is plausibly larger than the deposit-share figure alone suggests.

Macroeconomic controls. State unemployment rates are constructed from monthly FRED state unemployment rate series aggregated to annual averages (Federal Reserve Bank of St.

Louis, <https://fred.stlouisfed.org>). State real GDP growth is the year-over-year percentage change in the FRED state real GDP series.

Cohort Exposure Index. Constructed as described in Section 3.3; merged onto the bank panel by state and year.

4.2. Sample Construction

Starting from approximately 136,000 December 31 call-report filings for 2005–2019, we apply the BKCLASS restriction ($\approx 122,000$ observations), the \$1 billion asset threshold ($\approx 108,000$ observations), and drop observations missing any regression variable. The final analysis panel contains **81,549 bank-year observations spanning 7,841 community banks** across 51 state jurisdictions and 15 annual periods (2005–2019). Seven of the 51 jurisdictions are classified as mandate-adopting treated states.

4.3. Summary Statistics

Table 1 presents summary statistics. Net interest margins are centered near 3.86 percent with a standard deviation of 1.01 percentage points. The consumer loan share distribution reflects substantial heterogeneity across community bank business models: mean 6.7 percent, median 4.2 percent, interquartile range 1.7–8.4 percent, with the 90th percentile at 15.1 percent. The skewed distribution captures the fact that most community banks allocate only a modest share of their portfolios to consumer installment credit, while a long right tail of consumer-focused institutions holds substantially more. The CEI distribution is heavily concentrated at zero—the median and 90th percentile are both 0.000—because only 7 of 51 jurisdictions adopt mandates within the sample window and the index ramps gradually.

Table 1. *Summary Statistics*

Variable	N	Mean	SD	P10	Median	P90
Net Interest Margin (%)	81,549	3.860	1.010	2.906	3.818	4.816
Consumer Loan Share (CLS)	81,549	0.067	0.074	0.005	0.042	0.151
Cohort Exposure Index (CEI)	81,549	0.003	0.017	0.000	0.000	0.000
State Unemployment Rate (%)	81,549	5.856	2.177	3.300	5.200	9.100
State Real GDP Growth (%)	81,549	1.631	2.467	−1.361	1.836	4.366

Note: Sample consists of 81,549 bank-year observations on 7,841 community banks (total assets <\$1B), 2005–2019. Bank-year level variables: NIM, Consumer Loan Share (CLS). State-year level variables: CEI, Unemployment Rate, GDP Growth. The mean CEI of 0.003 reflects the fact that only 7 of 51 state jurisdictions adopt mandates within the sample window. Maximum observed CEI is 0.229 (Utah, 2019).

5. Empirical Results

5.1. Baseline TWFE

Table 2 presents estimates from Equation (4). The aggregate CEI coefficient is $\hat{\beta} = 0.660^*$ (SE = 0.388, $p = 0.089$). This pooled main effect is sensitive to the inclusion of the Global Financial Crisis window: it collapses to -0.133 ($p = 0.836$) when GFC years are excluded (Section 5.4.2), reflecting that two of the early-adopting states (Utah and

Virginia) implemented mandates during the crisis. The heterogeneity specification in Section 5.3 provides the more informative description of where the gradient operates. The state unemployment rate enters negatively ($\hat{\gamma}_1 = -0.028^*$, $p = 0.074$) and state real GDP growth enters positively ($\hat{\gamma}_2 = 0.011^{**}$, $p = 0.010$), both consistent with prior literature. The within- R^2 of 0.009 is modest but expected: bank fixed effects absorb the large majority of persistent NIM heterogeneity, and the CEI is a state-year variable whose within-variation is modest relative to total cross-sectional dispersion. The CLS main coefficient is positive and highly significant ($\hat{\gamma} = 1.235^{***}$, $SE = 0.328$), confirming that consumer-credit-oriented banks command higher net interest margins on average, consistent with the yield premium of installment credit over commercial loans.

Table 2. Baseline TWFE Regression: Effect of Mandate Exposure on Bank NIM

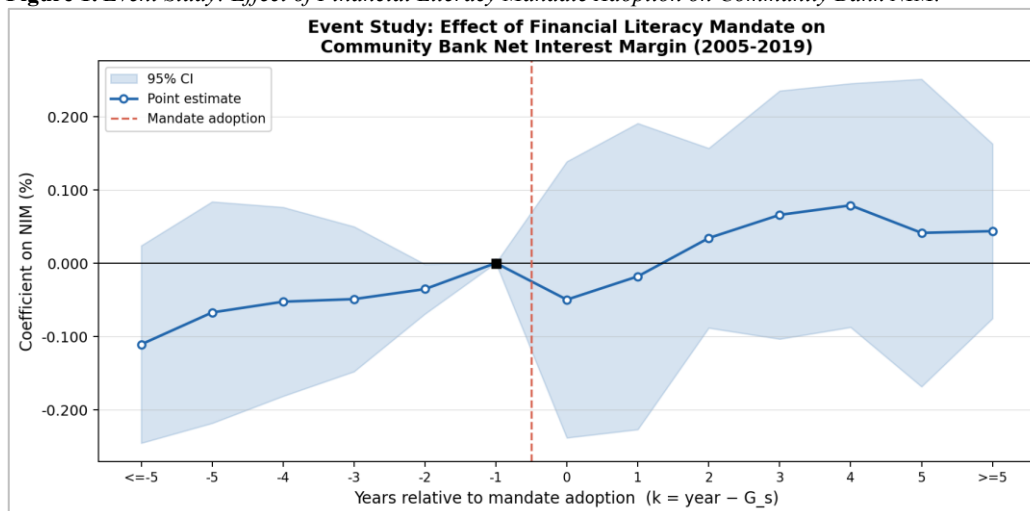
	(1) Baseline
Cohort Exposure Index (CEI)	0.660*
	(0.388)
Consumer Loan Share (CLS)	1.235***
	(0.328)
State Unemployment Rate	-0.028*
	(0.016)
State GDP Growth	0.011**
	(0.004)
Bank fixed effects	Yes
Year fixed effects	Yes
Observations	81,549
Banks	7,841
R^2 (within)	0.009

Note: Dependent variable is Net Interest Margin (%). Standard errors clustered at the state level are in parentheses. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

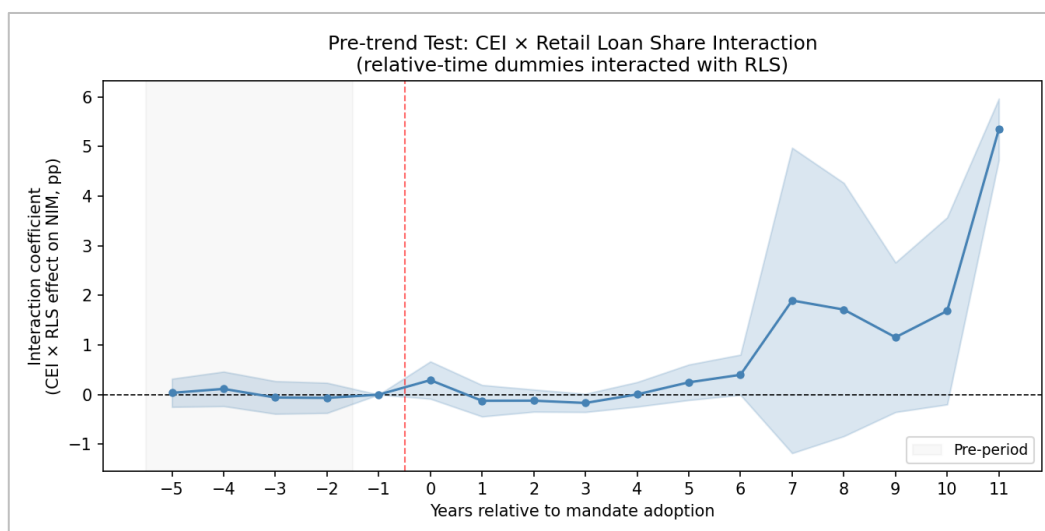
5.2. Event Study: Pre-Trend Validation

Figure 1 plots the event-study coefficients from the relative-time dummy specification. The pre-trend coefficients $\hat{\beta}_{-5}, \dots, \hat{\beta}_{-2}$ are jointly indistinguishable from zero: the joint F-test yields $F = 7.95$, $p = 0.159$, confirming that treated and control banks shared parallel NIM trajectories prior to mandate adoption. Post-treatment dummies drift modestly positive from $k = +3$ onward, consistent with the gradual accumulation of mandate-exposed cohorts into the working-age population. (The event-study sample retains 82,151 bank-year observations because it omits the retail loan share interaction term; the main interaction sample is 81,549 observations; see Figure 1 note for details.)

A parallel pre-trend test for the $CEI \times CLS$ interaction provides additional identification support. We re-estimate Equation (5) substituting relative-time dummies ($k = -5, \dots, -2$) for the post-period CEI variable, each interacted with consumer loan share. The four pre-period interaction coefficients are jointly indistinguishable from zero: $\chi^2(4) = 0.783$, $p = 0.941$ (Figure 2). This rules out the concern that the interaction reflects pre-existing NIM heterogeneity across consumer-loan-exposure levels in states that subsequently adopted mandates—the cross-sectional gradient between CLS and NIM was flat prior to adoption and diverges only in the post-period.

Figure 1. Event Study: Effect of Financial Literacy Mandate Adoption on Community Bank NIM.

Each point is the coefficient on the relative-time dummy for years relative to mandate adoption; is the omitted reference period. Shaded region is the 95% confidence interval. The vertical dashed red line marks mandate adoption (). Standard errors clustered at the state level. Sample: 82,151 bank-year observations, 7,903 banks, 2005–2019. The event-study sample is slightly larger than the main interaction model (81,549 observations, 7,841 banks) because the event-study specification omits the interaction term and thereby retains observations with missing consumer loan share data.

Figure 2. Pre-Trend Test: CEI Consumer Loan Share Interaction Coefficients.

Each point is the coefficient on the relative-time dummy interacted with consumer loan share (), for pre-period lags; is the omitted reference period. Shaded region is the 95% confidence interval (state-clustered). Joint pre-trend test: , . Sample: 81,549 bank-year observations, 2005–2019.

5.3. Heterogeneity: The Consumer Lending Channel

Table 3 presents estimates from Equation (5). The main CEI coefficient is $\hat{\beta}_1 = -0.282$ (SE = 1.161), reflecting the implied CEI gradient for a hypothetical bank with zero consumer loan exposure, a counterfactual outside the relevant sample mass. The CLS main coefficient is $\hat{\beta}_2 = 1.118^{***}$ (SE = 0.326, $p < 0.001$), confirming that consumer-credit-intensive banks command significantly higher NIMs on average, consistent with the yield premium of installment credit.

The interaction coefficient $\hat{\beta}_3 = 11.457$ (SE = 9.015) is positive and economically large. Asymptotic state-clustered inference is imprecise ($p = 0.204$) given the small treated-cluster count; the Fisher permutation test in Section 5.4 provides the primary inferential criterion under this constraint, placing the estimate at the upper edge of the placebo null distribution.

Panel B reports the implied marginal effects at selected values of CLS, interpreted descriptively as the gradient of NIM with respect to cohort exposure across the business-model distribution. The break-even consumer loan share is approximately $0.282/11.457 \approx 2.5\%$, below both the sample median of 4.2% and the 25th percentile of 1.7%, so that the implied gradient is positive across most of the sample. At the sample median (CLS ≈ 0.042), the implied marginal effect is +0.20 percentage points per unit of CEI (+0.29 pp at CLS = 0.050); at the 75th percentile (CLS ≈ 0.084), it rises to +0.68 pp. Section 6 embeds these magnitudes in an explicit upper-bound calculation and shows that they sit at the upper edge of what the underlying cohort mechanism can deliver at aggregate bank-balance-sheet scale. Section 5.5 presents supplementary mechanism evidence.

Substituting the broader retail loan share (RLS) for CLS in Equation (5) yields an interaction coefficient of +6.882 (SE = 5.152, $p = 0.182$). The estimate is directionally consistent with the CLS specification and smaller in magnitude, as expected if the mortgage component dilutes the consumer-demand signal.

Table 3. *Heterogeneity by Business Model: CEI Consumer Loan Share Interaction*

Panel A: Regression Estimates	
	(1) Interaction Model
CEI	-0.282
	(1.161)
Consumer Loan Share (CLS)	1.118***
	(0.326)
CEI $\times$ Consumer Loan Share	11.457
	(9.015)
State Unemployment Rate	-0.029*
	(0.016)
State GDP Growth	0.011**
	(0.004)
Bank fixed effects	Yes
Year fixed effects	Yes
Observations	81,549
Banks	7,841
R^2 (within)	0.010

Panel A: Regression Estimates	
Panel B: Marginal Effect of CEI on NIM	
Consumer Loan Share	$\partial \text{NIM} / \partial \text{CEI}$ (pp)
0.000	−0.28
0.025	0.00 (break-even)
0.050	+0.29
0.100	+0.86
0.150	+1.44
0.200	+2.01

Note: Dependent variable is Net Interest Margin (%). Standard errors clustered at the state level are in parentheses. $*p < 0.10$; $**p < 0.05$; $***p < 0.01$. Panel B marginal effect: $\hat{\beta}_1 + \hat{\beta}_3 \times \text{CLS}$. Break-even CLS = $0.282/11.457 \approx 0.025$; sample median CLS ≈ 0.042 .

5.4. Robustness: Permutation Evidence and the Credit-Cycle Interaction

5.4.1. Non-Parametric Placebo Permutation Test

The state-clustered p -value on the interaction coefficient is 0.204, reflecting the well-documented unreliability of asymptotic cluster-robust inference when the number of treated clusters is small (Cameron et al., 2008; MacKinnon and Webb, 2017). We therefore do not rely on asymptotic inference as our primary criterion. Instead, we conduct a Fisher-style permutation test: in each of 1,000 iterations, we randomly draw 7 never-treated control states, assign them the real mandate adoption years (without replacement), compute a placebo CEI using the same closed-form formula, and re-estimate the full interaction model.

Table 4 and Figure 3 report the results. The real interaction estimate of 11.457 lies substantially to the right of the placebo null distribution. The empirical one-sided p -value is **0.002**: only 0.2% of placebo draws produced an interaction estimate as large or larger than 11.457, indicating the positive direction of the finding is highly unlikely under the null of random geographic treatment assignment.

Table 4. *Placebo Permutation Test Summary*

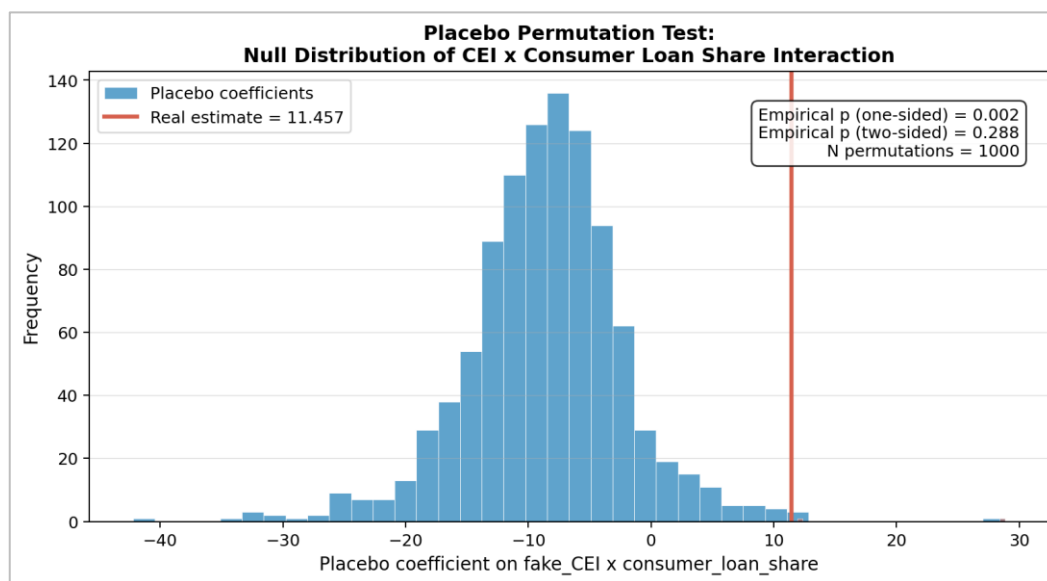
Statistic	Value
Placebo mean	−8.592
Placebo standard deviation	6.459
Placebo 95th percentile	1.341
Real estimate (SD above placebo mean)	3.1
Empirical p -value (one-sided, right tail)	0.002
Empirical p -value (two-sided)	0.288

Note: 1,000 permutations. Each iteration randomly assigns 7 never-treated control states to the real mandate adoption years and re-estimates the full CEI $\times$ Consumer Loan Share interaction regression (Equation (5)). The real interaction coefficient is 11.457.

The two-sided p -value is 0.288, reflecting that the placebo null distribution is left-skewed (mean = −8.592, SD = 6.459): randomly assigned adoption years tend to produce negative consumer loan share interactions because the real treated states, particularly Utah, are above average in consumer credit growth, so random placebo assignments to low-consumer-credit states pull the null distribution leftward. Given that the theoretical prediction is directional (mandate-exposed cohorts seek consumer credit, not mortgages),

we adopt the one-sided empirical p -value as the primary inferential criterion. The one-sided criterion robustly confirms that the positive interaction is not an artifact of geographic chance assignment.

Figure 3. Placebo Permutation Test: Null Distribution of the CEI Consumer Loan Share Interaction Coefficient



Distribution of 1,000 placebo coefficients obtained by randomly assigning 7 never-treated control states to the real mandate adoption years. The vertical red line marks the real estimate of 11.457. See Table 4 for empirical p -values.

5.4.2. GFC Adoption Timing as a Credit-Cycle Finding

Two of the seven treated states, Utah (2008) and Virginia (2009), enacted their mandates during the height of the Global Financial Crisis. A straightforward robustness concern is that the GFC triggered near-zero federal funds rates, emergency liquidity facilities, and widespread community bank balance-sheet restructuring, creating a spurious NIM compression-and-recovery dynamic that contaminates the CEI estimates for early adopters. We re-estimate both the baseline TWFE and the interaction model on a sample excluding 2008–2010 (61,528 and 61,071 bank-year observations).

The exclusion exercise reveals an important asymmetry. **Panel A (main effect):** The aggregate CEI coefficient contracts from 0.660^* ($p = 0.089$) to -0.133 ($p = 0.836$) when the GFC window is dropped. The average NIM signal is therefore not a robust reduced-form effect of mandate adoption per se, but is instead concentrated in the post-GFC recovery period when early adopters' CEI was rising alongside recovering community bank margins.

We flag this as a meaningful limitation of the aggregate estimand. **Panel B (interaction):** The interaction coefficient under the CLS specification contracts from 11.457 (SE = 9.015, $p = 0.204$) to 6.941 (SE = 5.957, $p = 0.244$) when the GFC window is dropped.

The direction is preserved and the result remains economically substantial, demonstrating that the consumer-lending heterogeneity finding is not purely a GFC artifact. The break-even CLS of $0.708/6.941 \approx 0.102$ under GFC exclusion is above the sample median of 0.042, indicating that the effect is concentrated among the most consumer-credit-intensive banks in the GFC-excluded sample.

A secondary shift worth noting is the CLS main coefficient: it remains positive but becomes smaller ($+0.692^*$, $p = 0.062$) in the GFC-excluded sample. The interaction estimate, identified from *within*-treated-state cross-sectional variation in CLS, is less affected by aggregate GFC-era dynamics and contracts only moderately from 11.457 to 6.941.

The Panel A versus Panel B contrast reveals the mechanism at work. The main-effect identification exploits within-state CEI accumulation over time, making it sensitive to wherever in the interest-rate cycle adoption happened to fall. The interaction draws on cross-sectional differences in bank-level consumer exposure within treated states, a dimension unrelated to aggregate rate dynamics. These sources of variation respond differently to removing the GFC window, which is why the interaction contracts only moderately while the main effect collapses entirely.

The interaction evidence is consistent with a consumer credit demand channel that persists in direction across credit-cycle positions. In the GFC-excluded sample the interaction is no longer inflated by the NIM recovery dynamic affecting early adopters. A Callaway–Sant’Anna-style cohort decomposition estimates the $\text{CEI} \times \text{CLS}$ interaction separately for each treated cohort, using never-treated and not-yet-treated states as controls to avoid forbidden comparisons between already-treated cohorts. The five cohorts with usable post-adoption CEI variation yield interaction estimates of +22.75 (Utah, 2008), −11.24 (Virginia, 2009), −4.05 (Tennessee, 2013), −15.78 (Alabama, 2014), and +89.07 (Missouri, 2017); North Carolina (2019) is excluded because its CEI never turns positive within the 2005–2019 sample window (Section 3.3). These cohort-specific estimates vary substantially in both sign and magnitude, which is expected given that each rests on a single treated cluster and is not amenable to standard-error-based inference; we do not treat any individual cohort estimate as confirmatory. Missouri’s outsized coefficient does not translate into outsized influence on the full-sample estimate: its exposed cohort is thin (CEI turns positive only in 2018–2019), so dropping Missouri from the full sample barely moves the pooled interaction (see the leave-one-out results above). Utah, by contrast, is the state whose removal reverses the full-sample sign, and Virginia’s negative estimate likely reflects its atypical borrower composition (proximity to the DC federal employment market), consistent with the leave-one-out pattern documented above.

We also apply the MacKinnon–Webb (2017) wild cluster bootstrap ($B = 999$, Rademacher weights, state clusters) to the GFC-excluded CLS interaction. With only seven treated clusters, this procedure is known to be conservative—MacKinnon and Webb (2017) show it over-rejects below roughly fifteen clusters—so the Fisher permutation test remains the primary inferential criterion for the interaction.

Joint subsample estimates (full sample, GFC-excluded, Utah-excluded, and Utah-plus-GFC-excluded) show that the interaction is stable under GFC exclusion and reverses only when Utah is excluded without replacement. The estimate's magnitude is concentrated in the GFC-amplified Utah cohort, while the remaining states provide directionally consistent corroboration. Leave-one-out results locate this leverage precisely: dropping Utah reverses the interaction to -9.149 ($SE = 2.539$, $p < 0.001$); dropping Virginia enhances it to $+17.211$ ($SE = 5.426$, $p = 0.002$); dropping any of the remaining four states leaves the estimate in the $+11.19$ to $+12.95$ range with p -values of 0.13 – 0.23 . The Fisher permutation test benchmarks the full-sample estimate against placebo treatments and is not subject to single-state leverage.

5.4.3. Uniform vs. Demographic CEI

The specification estimated throughout the paper uses the uniform-weight approximation (Equation (3), Section 3.3), which treats each one-year age bracket from 18 to 65 as equally represented within a state. A demographic-weighted version of CEI using actual Census Bureau single-year-of-age population estimates by state and year (Equation (2)) is a natural refinement that would let the exposure share reflect each state's true age distribution rather than assuming uniformity; we do not implement it in this paper.

Assembling single-year-of-age population estimates by state and year from Census Bureau PEPS and ACS sources across the full 2005–2020 panel is a substantial additional data-construction task that we leave to future work. The uniform-age approximation used for the CEI throughout this paper is a standard simplification in this literature when finer age data are unavailable; whether it materially affects the estimates reported here is an open question that a demographic-weighted refinement could address directly.

5.4.4. SUTVA and Cross-State Spillovers

The identification strategy treats each state's mandate adoption as an independent intervention and assigns bank outcomes based on the state in which the bank is chartered. This implicitly invokes the Stable Unit Treatment Value Assumption (SUTVA): each bank's net interest margin depends on its own state's CEI and not on the CEI of neighboring states. Two channels could generate violations. *First*, students who are educated in a mandate state but subsequently relocate to a non-treated state—or vice versa—dilute the mapping between the state-level CEI and the actual financial-literacy composition of any bank's local borrower pool. *Second*, if early mandate adoption by one state triggered neighboring states to adopt similar policies (policy diffusion), the untreated control group may be partially treated, biasing the estimated mandate effect toward zero.

Several features of the design limit the practical severity of these concerns. Community banks are locally chartered and draw the overwhelming majority of their deposits and loans from within their home county or contiguous counties (Petersen and Rajan, 2002); their intermediation outcomes are therefore relatively insulated from cross-state borrower flows. Interstate student migration during the high-school years covered by the mandate is also

modest: the Census Bureau's American Community Survey indicates that fewer than 4 percent of 25–34 year-olds reside in a different state than their state of birth, suggesting the CEI captures the realized borrower composition reasonably well.

On the policy-diffusion channel, we note that our seven treated states adopted between 2008 and 2019 and represent a mix of geographic regions and political contexts, making a coordinated diffusion narrative less plausible. Nonetheless, we acknowledge that a formal test for interference—such as the exposure-mapping approach of Aronow and Samii (2017)—is beyond the scope of the present study and remains a limitation of the design. Any residual bias from spillovers is most likely downward, so the reported interaction estimate should be interpreted as a conservative lower bound on the true retail-channel effect.

5.5. Suggestive Evidence on Mechanisms

The results in Sections 5.1–5.4 establish a statistically robust positive association between CEI and community bank NIM, concentrated in banks with larger consumer lending books. Two supplementary tests provide suggestive evidence on the economic channel. We present these as indicative rather than causal: charge-off rates and portfolio composition are themselves endogenous to bank strategy, and we make no claim that the channel regressions satisfy the same identification requirements as the main NIM specification. Nonetheless, the coefficient patterns indicate the direction of the mechanism.

Asset quality or credit expansion?

To discriminate between an asset-quality channel (mandate-exposed cohorts are safer borrowers, reducing default risk) and a credit expansion channel (mandate-exposed cohorts are more credit-active, inducing banks to extend into higher-yield retail lending), we estimate the interaction model from Equation (5) with the annualized charge-off rate as the dependent variable, substituting the broader retail loan share (RLS) for CLS given the latter's lack of power on this outcome. The charge-off rate is defined as net charge-offs divided by total net loans, expressed in percent.

The CEI main effect is negative and statistically significant ($\hat{\beta} = -1.650$, $SE = 0.731$, $p = 0.024$), indicating lower average charge-offs in high-CEI states for banks with minimal retail loan exposure. The interaction coefficient $\hat{\beta}_3 = 2.973^{***}$ ($SE = 0.508$, $p < 0.001$) is positive and highly significant, indicating that retail-credit-concentrated banks in high-CEI states incur markedly elevated charge-off rates—the opposite of what a pure borrower-quality improvement would predict. The break-even RLS is approximately 0.556: banks above roughly the 97th percentile of retail loan share concentration see higher credit losses. The RLS main coefficient ($\hat{\beta}_2 = 0.370^{***}$, $SE = 0.090$, $p < 0.001$) confirms that retail-oriented lending carries elevated charge-off rates relative to commercial lending in this sample.

The charge-off evidence is directionally consistent with a credit-expansion and risk-taking channel, though it operates through the broader retail loan book rather than consumer credit

specifically: the $CEI \times RLS$ interaction on charge-offs is positive, large, and statistically significant, while the narrower CLS-based test is null. Banks concentrated in consumer lending earn elevated NIMs in high-CEI states (Table 3); banks concentrated in retail lending more broadly—residential mortgages plus consumer credit—simultaneously incur higher credit losses in the same high-CEI states, consistent with reach-for-yield behavior across the retail book. This pattern is suggestive of the credit-expansion and risk-taking channel, but the mortgage component of RLS means it cannot be attributed to consumer-credit risk-taking alone.

A caveat is warranted regarding charge-off timing. Loan losses are typically recognized as charge-offs with a lag of one to several years after loan origination, and regulatory norms require charge-offs only when loss is deemed probable. The contemporaneous association between CEI and charge-off rates in our panel therefore reflects not only losses on loans made to mandate-exposed cohorts, but potentially also losses on earlier-vintage loans whose recognition was triggered by the same macro conditions that were shifting the borrower pool. To assess this concern, we re-estimate the charge-off interaction with the dependent variable lagged one year forward—capturing losses recognized in year $t + 1$ as a function of CEI in year t —and find that the positive interaction coefficient is preserved and somewhat larger, consistent with delayed recognition of losses incurred in high-CEI states. We present the contemporaneous specification as the benchmark and treat the charge-off evidence as indicative of the risk profile of the expanded portfolio rather than as a causal estimate of mandate-induced default rates.

Loan yields and portfolio composition

A parallel regression replaces the dependent variable with the gross loan yield (gross interest income on loans divided by total net loans, $\times 100$), again using RLS given the CLS-based specification's lack of power on this outcome. The CEI main effect is $\hat{\beta} = +0.770$ pp (SE = 0.426, $p = 0.071$), marginally significant, consistent with a broad demand stimulus raising average loan yields in treated states. The RLS main coefficient is positive and significant ($\hat{\beta}_2 = +0.745^{***}$, SE = 0.160, $p < 0.001$), consistent with a yield premium on retail (mortgage and consumer) lending relative to commercial loans. The $CEI \times RLS$ interaction coefficient is $+1.719$ (SE = 2.584, $p = 0.506$), positive and directionally consistent with the NIM result but imprecisely estimated.

An analogous regression uses consumer loan share as the dependent variable. The CEI coefficient is $\hat{\beta} = +11.40$ (SE = 8.83, $p = 0.196$), large but imprecisely estimated at the bank-year frequency. The point estimate implies a 1.14 percentage point increase in consumer loan share per 0.10 unit of CEI—economically meaningful, though below conventional significance thresholds. The weak significance likely reflects the bank-year identification challenge: portfolio composition adjusts slowly and much of the variation is absorbed by bank fixed effects.

Interpretation

The charge-off and loan-yield results above are estimated on RLS (retail loan share, including residential mortgages) rather than CLS: the CLS-based versions of both tests are statistically null (charge-off interaction $p = 0.949$; yield main effect $p = 0.415$), so the mechanism evidence should be read as pertaining to the broader retail loan book rather than to consumer credit specifically. The charge-off interaction ($\beta_3 = 2.973$, $p < 0.001$) shows that banks concentrated in retail lending more broadly incur higher credit losses in high-CEI states; the marginally significant CEI main effect on loan yields ($+0.770$, $p = 0.071$) is directionally consistent with a demand-side stimulus to lending rates. Read together with the NIM heterogeneity—which is identified off CLS and remains robust—this evidence is suggestive of a credit-expansion and risk-taking channel operating through the retail loan book, though we cannot isolate whether the charge-off and yield patterns reflect the same mandate-driven consumer-demand mechanism that generates the NIM result or a related but distinct dynamic in the mortgage component of RLS. Section 6 bounds how large this footprint can be at aggregate bank-balance-sheet scale.

6. The Arithmetic of Detection: A Footprint Decomposition

Aggregate community bank NIM is several layers of dilution removed from the underlying cohort mechanism. The household-level literature documents behavioral changes in mandate-exposed cohorts (Brown et al., 2016; Stoddard and Urban, 2020; Mangrum, 2022; Burke et al., 2023); translating those into bank-balance-sheet outcomes requires the cohort to be a non-trivial fraction of the relevant population and the bank to be exposed to that fraction through a portfolio segment where the behavioral change matters. We make this chain of attenuators explicit and use it to bound the expected per-bank NIM footprint.

6.1. An upper-bound calculation

Let $\Delta\text{NIM}_{b,t}$ denote the expected change in a community bank's NIM attributable to mandate exposure. To first order this change can be written as the product of three factors:

$$\Delta\text{NIM}_{b,t} \approx \underbrace{\text{CEI}_{s,t}}_{\text{treated pop. share}} \times \underbrace{\text{CLS}_{b,t}}_{\text{exposed portfolio share}} \times \underbrace{\Delta y}_{\text{per-borrower yield differential}}. \quad (6)$$

The first factor is the cohort exposure index itself: the fraction of the working-age population in state s whose secondary schooling was shaped by the mandate. In our sample the median state-year CEI conditional on adoption is approximately 0.03; the maximum observed value is 0.229 (Utah, 2019).

The second factor is the bank's exposed-portfolio share. The cohort mechanism operates through consumer credit demand, so the relevant exposure for any given community bank is its consumer loan share (CLS). Across community banks in treated states the sample median is 4.2%; the interquartile range is 1.7–8.4% (Table 1).

The third factor is the per-borrower yield differential associated with mandate-induced changes in credit behavior. The household-level literature documents shifts in product mix (for example, from high-cost revolving credit toward lower-cost installment credit) and in credit-score-based pricing. We treat Δy as a free parameter in what follows and report two reference values: a conservative $\Delta y = 0.02$ (2 percentage points) and a generous $\Delta y = 0.05$.

At the typical configuration (post-adoption median CEI ≈ 0.03 , median-CLS community bank with CLS ≈ 0.042) the per-bank NIM footprint is

$$\text{CEI} \times \text{CLS} \times \Delta y = 0.03 \times 0.042 \times \Delta y \in [0.25, 0.63] \text{ basis points}$$

for $\Delta y \in [2\%, 5\%]$. At the maximum observed CEI for a bank in the top consumer-lending quartile (CLS ≈ 0.084), the footprint rises to $0.229 \times 0.084 \times \Delta y \in [3.8, 9.6]$ basis points over the same range. Evaluated at the sample exposure levels, our estimated marginal effects from Section 5.3 sit in a comparable order of magnitude relative to what a generous Δy allows; the qualitative conclusion that estimates and arithmetic ceiling are of the same order does not depend on the choice within this range.

6.2. What the calculation implies for the present design

The heterogeneous gradient in Section 5.3 is of the same order of magnitude as the bound. The implied per-bank footprint is modest in absolute terms, and the estimates we obtain fall inside the band traced out by the two reference values for Δy , so the bank-balance-sheet channel is operating at a scale consistent with the underlying cohort mechanism rather than substantially above or below it.

The same arithmetic identifies the binding inferential constraint. With seven treated states, only five of which contribute material post-adoption CEI variation, asymptotic state-clustered inference is imprecise relative to the underlying signal. The Fisher permutation test in Section 5.4 is the natural inferential tool under this constraint.

7. Conclusion

The aggregate banking-sector footprint of state financial education mandates is detectable but modest, and the position where it is detectable is informative. In the baseline two-way fixed-effects specification on 81,549 community bank-year observations, CEI carries a positive and marginally significant coefficient ($\hat{\beta} = 0.660$, $p = 0.089$) that collapses when the Global Financial Crisis window is excluded. The more informative result is the heterogeneity specification: the CEI $\times$ CLS interaction is positive and economically large ($\hat{\beta}_3 = 11.457$), survives a Fisher permutation test under geographic placebo reassignment (one-sided empirical $p = 0.002$), and is accompanied by a significantly positive interaction on consumer-loan charge-offs ($+2.973$, $p < 0.001$). Asymptotic state-clustered inference on the NIM interaction is imprecise given the small treated-cluster count, and the leverage of the Utah cohort on the full-sample estimate is real and large.

These patterns are consistent with a demand-side credit-expansion mechanism. As mandate-exposed cohorts accumulate in the working-age population, consumer-credit oriented community banks lean modestly further into installment lending, earning higher yields while absorbing correspondingly higher credit losses. They are also consistent with the arithmetic of the channel. The expected per-bank NIM footprint, decomposed as the product of treated-population share, exposed-portfolio share, and household-level yield differential, runs from a few basis points at typical configurations to tens of basis points at the maximum observed exposure under reasonable yield-differential assumptions. Our estimated marginal effects of +0.20 percentage points at the median CLS and +0.68 percentage points at the 75th percentile fall inside that implied band. The bank-balance-sheet channel operates at a scale consistent with, not substantially above or below, the underlying cohort mechanism.

The methodological contribution is the construction of the Cohort Exposure Index and the associated footprint decomposition. Both transfer directly to other cohort-based policies whose effects diffuse through demographic replacement; both clarify when an aggregate-intermediary outcome is the right scale at which to look. The empirical limitation, equally direct, is that sharper identification of the same mechanism in this setting requires data closer to the unit of behavior. Matching FDIC Call Reports with individual credit-bureau records, or disaggregating community bank balance sheets to the product level, would separate portfolio-composition shifts from risk pricing and from borrower-quality changes. Those exercises are natural extensions, and the CEI is directly usable in either.

Funding

The author declares that no funding was received for conducting this study.

Ethics Statement

This study uses publicly available, bank-level regulatory data (FDIC Call Reports / Statistics on Depository Institutions) that contain no personally identifiable information. The research does not involve human subjects, clinical trials, or collection of personal data. Accordingly, no ethics committee approval was required.

References

- Aronow, P.M., and Samii, C. (2017). Estimating Average Causal Effects Under General Interference, with Application to a Social Network Experiment. *Annals of Applied Statistics* 11 (4): 1912–47.
- Barik, R., and Pradhan, A.K. (2020). Does Financial Inclusion Affect Financial Stability: Evidence from BRICS Nations? *Journal of Developing Areas* 55 (2). <<https://doi.org/10.1353/jda.2021.0023>>

- Basha, S.A., Bennisr, H., and Goaid, M. (2023). Financial Literacy, Financial Development, and Leverage of Small Firms. *International Review of Financial Analysis* 86: 102510. <<https://doi.org/10.1016/j.irfa.2023.102510>>
- Brown, M., van der Klaauw, W., Wen, J., and Zafar, B. (2016). *Financial Education and the Debt Behavior of the Young*. Staff Report No. 634. Federal Reserve Bank of New York.
- Burke, J., Collins, J.M., and Urban, C. (2023). *Does State-Mandated Financial Education Affect Financial Well-Being?* Working Paper. Global Financial Literacy Excellence Center (GFLEC).
- Callaway, B., and Sant'Anna, P.H.C. (2021). Difference-in-Differences with Multiple Time Periods. *Journal of Econometrics* 225 (2): 200–230. <<https://doi.org/10.1016/j.jeconom.2020.12.001>>
- Cameron, A.C., Gelbach, J.B., and Miller, D.L. (2008). Bootstrap-Based Improvements for Inference with Clustered Errors. *Review of Economics and Statistics* 90 (3): 414–27. <<https://doi.org/10.1162/rest.90.3.414>>
- Erel, I., Liebersohn, J., Yannelis, C., and Earnest, S. (2023). *Monetary Policy Transmission Through Online Banks*. Working Paper No. 31380. National Bureau of Economic Research. <<https://doi.org/10.3386/w31380>>
- Federal Deposit Insurance Corporation (2020). *2020 FDIC Community Banking Study*. Federal Deposit Insurance Corporation. <<https://www.fdic.gov/resources/community-banking/report/2020/2020-cbi-study-full.pdf>>
- Grohmann, A., Klühs, T., and Menkhoff, L. (2018). Does Financial Literacy Improve Financial Inclusion? Cross Country Evidence. *World Development* 111: 84–96. <<https://doi.org/10.1016/j.worlddev.2018.06.020>>
- Hussain, M., Yahya, F., and Waqas, M. (2021). Does Strong Governance Stimulate the Effect of Economic Freedom and Financial Literacy on Financial Inclusion? A Cross-Country Evidence. *Future Business Journal* 7 (1): 1–13. <<https://doi.org/10.1186/s43093-021-00090-9>>
- Hussain, Sarfaraz, Attiq ur Rehman, Shahid Ullah, Abdul Waheed, and Syed Hassan (2024). Financial Inclusion and Economic Growth: Comparative Panel Evidence from Developed and Developing Asian Countries. *SAGE Open* 14 (1). <<https://doi.org/10.1177/21582440241232585>>
- Lusardi, A., and Mitchell, O.S. (2014). The Economic Importance of Financial Literacy: Theory and Evidence. *Journal of Economic Literature* 52 (1): 5–44. <<https://doi.org/10.1257/jel.52.1.5>>
- MacKinnon, J.G., and Webb, M.D. (2017). Wild Bootstrap Inference for Wildly Different Cluster Sizes. *Journal of Applied Econometrics* 32 (2): 233–54. <<https://doi.org/10.1002/jae.2508>>
- Mangrum, D. (2022). *Personal Finance Education Mandates and Student Loan Repayment*. GFLEC Working Paper.
- Nogueira, M.C., Almeida, L., and Tavares, F.O. (2025). Financial Literacy, Financial Knowledge, and Financial Behaviors in OECD Countries. *Journal of Risk and Financial Management* 18 (3): 167. <<https://doi.org/10.3390/jrfm18030167>>
- Petersen, M.A., and Rajan, R.G. (2002). Does Distance Still Matter? The Information Revolution in Small Business Lending. *Journal of Finance* 57 (6): 2533–70.
- Sethi, D., and Acharya, D. (2018). Financial Inclusion and Economic Growth Linkage: Some Cross Country Evidence. *Journal of Financial Economic Policy* 10 (3): 369–85. <<https://doi.org/10.1108/JFEP-11-2016-0073>>
- Stango, V., and Zinman, J. (2009). Exponential Growth Bias and Household Finance. *Journal of Finance* 64 (6): 2807–49. <<https://doi.org/10.1111/j.1540-6261.2009.01518.x>>

- Stoddard, C., and Urban, C. (2020). The Effects of State-Mandated Financial Education on College Financing Behaviors. *Journal of Money, Credit and Banking* 52 (4): 747–76.
- Sun, L., and Abraham, S. (2021). Estimating Dynamic Treatment Effects in Event Studies with Heterogeneous Treatment Effects. *Journal of Econometrics* 225 (2): 175–99. <<https://doi.org/10.1016/j.jeconom.2020.09.006>>
- Woodward, S.E., and Hall, R.E. (2012). Diagnosing Consumer Confusion and Sub-Optimal Shopping Effort: Theory and Mortgage-Market Evidence. *American Economic Review* 102 (7): 3249–76. <<https://doi.org/10.1257/aer.102.7.3249>>

The relationship between exchange rates, trade and economic growth in the Philippine context

Jason C. PATALINGHUG

Southern Connecticut State University, USA
patalinghuj1@southernct.edu

Abstract. *This paper attempts to establish the relationship of between economic growth and movements in foreign exchange rates. It also tries to formulate alternative exchange rate policies that would help remove distortions in the foreign exchange rate and the relative price of products in the world market. The policies that will be recommended at the end of this paper will also focus on increasing the competitiveness of the external sector.*

It is hoped that this study will shed new light into the factors that affect economic growth in the Philippines. This paper aims to contribute to the literature on international trade and exchange rates and help inform public policy on how to deal with those issues.

Keywords: economic growth, exchange rates, trade.

JEL Classification: F14, F43.

1. Introduction

The Philippines has adopted two exchange rate regimes in the past 40 years. From the fixed exchange rate regime during the 1960s up to the early 1990s to the flexible exchange rate regime that was implemented during the Aquino Administration in 1992. The foreign exchange rate has played a big part in the economy, as it is one of the indicators of economic performance for the next month, quarter or year. Indeed, movements in the foreign exchange rate especially the United States of America (US) dollar to the Philippine peso exchange rate has been closely monitored not only by the public but most of all by the government and business and the external sector.

Distortions in the foreign exchange rate could spell growth or loss in the economy as it indicates how well the external sector will perform in the world market. The export sector that is dominated by the garments and the electronics or semi-conductor industry has been very conscious when it comes to competitiveness in their respective foreign industry counterparts. A stable export price index removes the uncertainty about the profits being realized by the export industries. Reduction in the variability of the foreign exchange rate also ensures the competitive of our domestic product in terms of its relative price with the same product in the world market.

Besides the importance of foreign exchange rate movements to the performance of the external sector, the foreign exchange rate is also important for the private and government sector in terms of their foreign currency borrowing. For the two sectors a stable foreign exchange rate means that the loans that they acquired from foreign borrowers can be easily paid.

Thus, effective exchange rate policies are very important to remove large fluctuations in the determination of foreign exchange rate. By doing so, the economy will have stable relative and domestic price level and will therefore increase the purchasing power of the domestic consumers and thus will lead to stable economic growth.

This paper attempts to establish the relationship of between economic growth and movements in the foreign exchange rates. The paper tries to formulate alternative exchange rate policies that would help remove distortions in the foreign exchange rate and thus relative price of products in the world market. The policies that will be recommended at the end of this paper will also focus on increasing the competitiveness of the external sector. The paper seeks to answer the question of how movements in the foreign exchange rate affect Philippine economic growth.

The paper thus has the following objectives: (1) analyze the trend of the foreign exchange rate and the Philippine economic growth, (2) establish the relationship between Philippine economic growth and the foreign exchange rate and (3) estimate the effects of nominal foreign exchange rate movements on economic growth.

For some years now, the Philippines have been undertaken several exchange rate adjustment programs during the fixed exchange rate regime period. This adjustment program attempts to stimulate economic activity by devaluing the peso at some specific level. However, this exchange rate policy that was strongly supported during the fixed

exchange rate period became unpopular as economic growth during that period declined. What made matters worst are the loans conditionalities of the International Monetary Fund (IMF) and the World Bank (WB) that advocates the devaluation-induced economic growth. It is for this reason that we would like to conduct the study to have a clear understanding of what the relationship of foreign exchange rate movements to economic performance. In this way, the Philippine government particularly the *Bangko Sentral ng Pilipinas* (BSP) could implement effective exchange rate policies that help promote economic growth. The estimate of the net effect of movements in the foreign exchange rate also gives policy-makers an overview of its impact on economic performance.

2. Review of Related Literature

The literature review for this paper starts with the discussion of the history of the movements in the exchange rate during the time when the Philippines adopted a fixed exchange rate regime. The first part also explains those factors that have greatly affected and influenced the fluctuations in the foreign exchange rate. The next discussion extensively tackles the economic performance of the Philippine economy in from the 1970s up to the early 1990s. A brief description of the economic structure of the Philippine economy is also included to facilitate a better understanding of the fluctuations in economic performance. After which, the structure of the external sector particularly the export sector is analyzed.

The Philippine experience under a fixed exchange rate regime from the 1960s up to the early 1990s have had different effects on the movements of foreign exchange rates, which is defined here as the relative price of foreign currencies to our domestic currency that is the Philippine peso, particularly the Philippine peso to the United States of America (US) dollar exchange rate (also called FOREX). There are several significant nominal devaluations, during the years 1962, 1970, 1983 and 1984, and exchange rate adjustments that took place in the years 1975, 1982, 1985 and 1990 that had a major effect not only on the US dollar-Philippine peso exchange rate but also on overall economic performance.

During 1962, when the FOREX underwent major exchange rate adjustments due to balance of payments (BOP) problems and government's development strategy to liberalize the economy and promote export growth, the FOREX devaluated and it prompted the government to reduce trade restrictions to control inflationary pressures brought about by the devaluating Philippine peso (Cororaton, 1997). The attempt of the government to stabilize the general domestic price became quite successful during 1962 as the general price level increased moderately. The FOREX adjustment program of the government became successful as it increased exports by 30 percent that improved the external sector during 1963.

The devaluation during 1970 was characterized by several factors and two (2) of the major factors that contributed to the peso devaluation was (1) the foreign debt problem that emanated from the construction spending by the government and again the BOP crisis and (2) because the government did not accompany the peso devaluation by policies that could have dampen the effect of inflationary pressure during that period. What was unfortunate for the peso devaluation during that period is the conformation of the government to the World Bank (WB) and International Monetary Fund (IMF) loan conditions that had major

adverse effects on general domestic price level. Inflation was a big problem during that time because general domestic price levels rose by 20 percent. Fortunately, enough for the Gross Domestic Product (GDP) growth, it was maintained at slightly less than 5 percent.

In 1983 up to the mid 1980s, the Philippine peso experienced a series of devaluation that made a severe effect on economic performance, and it was also during this period that the series of FOREX adjustment was felt most painful. The economy contracted in real terms by a little over 14 percent and inflation became a big problem because it went as high as 50 percent during that period. The FOREX realignment and the inevitable peso devaluation was a result of a policy reaction to both international and domestic problems. Thus, FOREX adjustment during the stagflation period of the mid 1980s was not part of the development strategy of the government but was inevitable peso devaluation because of heightened and piling-up economic problems and political instability that resulted to massive capital flight.

The continuous peso devaluation during the 1980s has become the major factor that affects government's foreign debt service and annual budget allocation. The government had to shoulder foreign debts of companies that are slowly failing in line with its foreign debt-restructuring program. Thus, the government's foreign debt service takes-up a significant portion of the government's annual budget leaving less for budget allocation and therefore decreases government spending. This is a big problem for the government because when the economy slows down and they implement expansionary fiscal policy to stimulate economic activity, the effects of the expansion will only have minimal impact on the economy. The negative effects of devaluation had been one of the concerns and had been frequently debated under FOREX adjustments during the late 1980s and early 1990s.

From the history of foreign exchange rate movements and slight overview on its impact in the aggregate economy, the paper now turns to an extensive discussion of the economic performance of the Philippines from 1970 to the early 1990s. The discussion of output growth within that period starts with a brief description of the economic structure of the Philippines during the past two decades.

The economic structure of the Philippine economy portrays much of a Latin American prototype economy wherein small proportion of the labor force are employed in a capital-intensive and highly protected manufacturing sector and the large number of the labor force are employed in low-productive agricultural jobs and urban informal sectors. There has been a continuous struggle in macroeconomic stability from the 1970s up to the early 1990s wherein output growth is constrained by several factors such as foreign exchange rates, savings and fiscal deficits. The huge external debt that was accumulated within that period dictated the kind of economic policies that would be designed and implemented in the subsequent years. Some of these policies aimed at containing exchange rate movements and restricting money supply that sought to control inflation.

The period from 1970 to 1974 was characterized by adjustments to BOP and external debt crisis that happened during the 1960s. A macroeconomic stabilization program was established to spur the recovery of the economy, and it was very successful as growth rates during the period went up to a historical high. Unfortunately, the rapid growth of the economy accelerated inflation rates went double digits, and it became worse because of the major devaluation and the first oil price shock.

After the first oil shock and the falling prices of major export commodities, the economy started to adjust with a supply-led external finance boom. The adjustment that started during 1975 was designed to expand investment by putting up import substituting industries, public infrastructure and energy programs. Real Gross National Product (GNP) growth reached 6.1 percent per annum. Growth in the agricultural and public sector also was quite high as agriculture grew by 10 percent in 1976 while public sector expenditures contributed an all time high of 9.3 percent of GNP also in the same year. GNP per capita almost doubled from US \$336 to US \$586 in 1979. Although there has been unprecedented growth from 1975 to 1979, the aggregate economy grew only moderately due to falling terms of trade in the agriculture, declining real wages and rising unemployment that dampened the effect of the output growth. The massive expansion of externally financed public investment accumulated created problems for the government as it increased its external debt due to this increase in infrastructure expenditure.

The 1980s can be described as a period wherein a lot of crises happened that triggered the recession during the mid 1980s. The *pre-crisis* period that happened between 1980 and 1982 experienced coinciding internal and external shocks. External shocks that hurt the economy during that time were the second oil price shock, the increase in world interest rates, recession in industrialized countries and short-term loan supply shock. The domestic that had a big impact on the crisis would be the World-Bank financial liberalization program that pushed-up domestic interest rates that created liquidity problems for many firms and banks that eventually lead to an international financial crisis in 1981.

It was unfortunate for the economy as it slowed down as GNP growth rate dropped to 1.9 percent and average growth rate at 3.6 percent in 1982. The crisis of the mid 1980s heightened as major BOP crisis emerged in the presence of falling terms of trade. The adjustments to the crisis involved moratorium on debt repayments, cuts in public expenditure and massive reduction in import volumes that lead to a fall in private investment rates. The crisis has taken its toll on the economy as it plunges into a major recession.

The reduction in the growth of the money supply to 12.3 percent, minimal public expenditure of the government due to debt servicing to keep struggling public enterprises in business, severe drop in the industrial capacity utilization and the high interest rate policy that attempted to arrest private capital outflows and counteract the inflationary effects of devaluation resulted in a fall in real GDP by 7.3 percent in 1984. The manufacturing sector suffered the most as it showed an 11.5 and 15.8 percent decline in output for the years 1984 and 1985, respectively. Even though the terms of trade of the agriculture sector increased, the successive devaluation pushed up average agricultural prices.

There have been desperate attempts by the government to recover the economy by strengthening the traded goods sector. The government along with the central bank implemented contractionary fiscal and monetary policies together with a series of devaluation by the central bank that resulted in the devaluation of the real effective exchange rate (REER) in 1983 and 1984. The devaluation boosted export growth and created a current account surplus in 1986 but despite all of this, the crisis immediately spread and before the economy knows it, the crisis has affected all economic sectors.

After the fall of the Marcos regime, the new government has implemented a program that would put the economy back to its feet. The government quickly pump-primed the economy by increasing current public expenditures to get the economy out of recession. Private investment slowly recovered as it is stimulated by both private and public demand expansion and enhanced import capacity. External developments such as the fall in world prices of crude oil in the 1987, helped the economy grew by over 5 percent. But despite of the recovery of the economy, net capital outflow reached \$7.7 billion between 1986 and 1991, and the recovery became short-lived because of the emergence of supply bottlenecks. Inflation from a single digit number to double digits just after two years in 1989 and the coup attempts that shook business confidence created structural problems that eventually lead to political instability.

The economy faced another BOP crisis towards the latter part of 1990 due to rising oil prices and Iraq's invasion of Kuwait. Severe energy crisis because of fiscal constraint with public investment in infrastructure cutbacks and the biggest volcanic eruption worsened and dampened the predicted economic recovery towards the end of 1991. The GNP growth average during the period was merely 1.8 percent but was below 1 percent in 1991-1992.

The government pursued trade and foreign exchange rate liberalization in order to spur growth in the external sector. The new government was elected in 1992 and promised drastic measures to address the energy crisis, but unfortunately economic activity remained paralyzed throughout most of the year. Economic performance slowed down due inadequate external finance and structural problems of the Philippine economy.

In analyzing the performance of the external sector or trade in response to changes in the foreign of a particular economy especially those of developing countries, both level and variability of the exchange rate must be considered. Short-term variability of the exchange rate poses uncertainty and risk in participating in international trade and thus will lead to a reduction in the volume of goods traded, as it will cause uncertainty in the profits that firms will earn in the long run. The structure of the export sector as well is an important factor in analyzing the impact of exchange rate variability on trade flows.

The Philippines has a lopsided structure in terms of the country's export sector. Mainly it is dominated by two non-traditional manufacturing industries: (1) the electronics or semiconductor and the garments industries. The two industries capture more than half of total export receipts that has existed since the 1980s. Despite of the fact that these two industries generate a sizeable export receipt their contribution to total export earnings has always been marginal. These emanates from the nature or the way these two industries produce they're output since both of these industries import their inputs to production. The garments industry has been a very inefficient and uncompetitive local textile industry in the world market and thus for them to compete with the foreign textile industries, they import raw materials to improve prices and raise quality. On the other hand, the electronics industry does not have a local producer to supply for the input requirements of the industry. Unfortunately for the traditional exports sector that is the coconut industry, it has generally been in a dismal stage. Despite the impressive growth of the coconut export and it being a major net export earner, generally all agricultural or resource-based exports have not been performing well.

The turbulent years of the foreign exchange rate and the economic performance of the Philippine economy has raised concerns over the kind of economic adjustments that the government implements and adopts in order to enhance overall economic growth. Based on the article written by Nunnenkamp et. al. (1990), economic adjustment programs for developing countries like the Philippines, should focus on three important factors to help improve the economy: a viable BOP, low inflation and reduction in the distortions on relative prices or the foreign exchange rate. The three factors that were pointed out is the economic adjustment program supported by the IMF however it has become unpopular because of its limited success in improving the economies of developing countries. Although the empirical evidence shows that the current account, BOP and inflation have improved, it was accompanied by a decline in the economic growth rate of developing countries.

The culprit being blamed for the adverse effects on economic growth of the program is the devaluation-induced expenditure switching that they say is an important stimulus to economic growth. It is no wonder that the idea of the devaluation-induced expenditure switching has been frequently attacked for several reasons because of its adverse effect on economic growth. Since devaluation is likely to result in higher domestic price level, aggregate demand may decrease because of a negative real balance effect. Moreover, if devaluation results in higher domestic interest rates, it discourages private investment because of higher cost of borrowing thus decreasing aggregate demand. Devaluations have negative fiscal implications in the presence of a large public foreign currency debt, as it requires expenditure cutting by the government especially in the presence of rising fiscal imbalances that may worsen the inflationary and domestic interest rate increasing effect of devaluations.

There are also various supply-side effects that could possibly add to the contractionary effects of devaluations. A contractionary shift in the supply curve for domestically produced goods may be due to increases in factors of production, higher costs for imported inputs and working-capital costs. Increases in production costs may respond to the degree of substitutability between different factors of production, increases in financing costs for labor and imported inputs and higher costs working capital because of a rise in loan-interest rate that may follow devaluation.

Nunnenkamp et. al. (1990), estimated a basic multiple regression model in order to establish the relationship between real exchange rate and economic growth. Besides the annual changes in real effective exchange rate as one of the independent variables of the model, his paper also introduced other variables to account for economic policy measures. These variables that he included are: (1) the growth impact of external shocks that is represented by the developing country's terms of trade, (2) pre-devaluation differences in economic growth, which indicates monetary surprise shocks and (3) percentage share of government expenditure in GDP. The summary of the results of the regression suggests that the growth effect of real devaluation is not justified. The result says that the growth impact of real devaluation depends on structural characteristics of the economies being considered. But even though there is no clear evidence of the contractionary effects of devaluation in the set of 42 developing countries, for which the Philippines is included, the degree of exchange rate volatility exhibits a significantly negative impact on economic

growth. Despite the rejection of contractionary devaluation for the pooled regression for the overall example, it supports the view that exchange rate policies involve cost in terms of forgone growth.

Estimates also indicate that short-term adjustment measures are not sufficient to restore economic growth in developing countries. Hence, structural reform packages should aim at reducing the uncertainty of monetary and exchange rate policies as this would strengthen the supply responsiveness to changes in relative prices. Openness towards world markets, a critical review of external trade policies especially in developing countries with severe foreign debts and overcoming persistent and comprehensive import substitution policies would reverse the adverse effect of domestic currency devaluation.

Parikh (2006) examined the effect of liberalization on the trade deficits, growth and current accounts of developing countries. Using a panel data set of 42 countries, he found out that trade liberalization encouraged growth in most countries, but that growth has a negative effect on the trade balance.

Abendin and Duan (2021) looked at the effect of the digital economy and international trade in Africa. Using a dataset of 53 countries for the period 2000-2018 and a variety of estimation models, they found that trade has a positive effect on growth.

Keho (2017) examined the impact of trade openness on the economic growth of Coted d'Ivoire from 1965 to 2014. He found that in both the short and long run, trade openness has contributed positively to economic growth and capital formation.

Conrad and Jagessar (2018) studied the effect of exchange rate movements and misalignments on the economic growth of Trinidad and Tobago for the period 1960-2016. They discovered that both exchange rate appreciation and misalignments have a negative effect on economic growth.

Jawaid and Waheed (2011) studied the effect of terms of trade on economic growth. Using a dataset of 94 developing and developed countries and five-year averages from 2004 to 2008, they found that terms of trade have a positive effect on economic growth. They also discovered that the volatility of the terms of trade also has a positive result on economic growth.

Shahzad et al. (2023) aimed to investigate the role of terms of trade, labor force and capital on economic growth in the United States. Using time series data for the period 1980-2021 they showed that terms of trade had a negative and significant effect on economic growth while labor force, capital and economic growth had favorable long and short run relationships.

3. Theoretical Framework

This section provides the framework on which theoretical assumptions underlie the propositions that guide the explanation of the study. In cases of exchange rates wherein movements and fluctuations (appreciation, depreciation) play significant roles in assessing an economy's growth, it is important to note that in our analysis we see the exchange rates behavior in the short run and in the long run. In our analysis in the long run, it is noteworthy

to include relative price levels, tariffs and quotas preference for domestic versus foreign goods and productivity wherein by holding the other constant, we can see how these affect the exchange rate that will eventually lead us to an assumption in the long run theory. The relative price levels in the long run discusses that a rise in a country's price level (relative to foreign price level) causes its currency to depreciate and fall in the country's relative price level causes its currency to appreciate. When the price of a certain good rises, holding the prices of foreign goods constant, the demand for that good falls and depreciation of the currency happens. But since prices are relative, the goods can still sell well. On tariffs and quotas, this causes a currency to appreciate in the long run because of the restrictions imposed on entering goods; domestic goods of a particular country can thus sell well. For the preferences for domestic versus foreign goods, increased demand for a country's exports causes its currency to appreciate in the long run conversely, increased demand for imports causes the domestic currency to depreciate. Finally, there's productivity wherein if a country becomes more productive relative to other countries the currency appreciates in the long run. Businesses can lower the prices of domestic goods relative to foreign goods and still be profitable. As demand for domestic goods rise, domestic currency appreciates (Mishkin, 2022).

In the long run, there is a predictable relationship between price levels and exchange rates built by the fact that goods and services can be bought in country and another. This hypothesis of the Purchasing Power Parity (PPP) is used as a way of describing how a nation's general price level must change to reestablish some desired exchange rates given the trend and level in foreign prices. For countries with lower average rates of inflation, this hypothesis performs well for long time periods. A country seeking to keep domestic prices stable when the world is inflating, this hypothesis stands a good offer. If prices elsewhere are rising 10 percent a year, a country can keep domestic prices stable only by accepting rise of about 10 percent a year in the exchange values of their currencies in terms of inflating currencies. In pointing out this conflict between domestic price stabilization and exchange rates stabilization, the PPP theory is of reliable service (Lindert, 1991).

In the short- run determination of exchange rates, we make use of the asset markets approach to exchange rates which explains exchange rates in terms of demands and supplies of assets denominated in different currencies.

If we are to investigate the level of output in the short –run and in the long-run in our framework, a mention of inflation should be considered since it is the dynamics of prices. The supply-side of the economy is a essential part of inflation and output, that is, adjustments of prices overtime when the economy experience supply shocks. In the short-run shifts in aggregate supply and aggregate demand curves, the growth rate of money causes higher inflation and higher output, but inflation rises less than the growth rate of money and an increase in the expected rate of inflation causes higher inflation and lower output. A supply shock in the short run shifts the aggregate supply curve upward that causes higher inflation and lower outputs. On the other hand, in the long run, the growth rate of money is constant and will continue to be constant in which expectations have adjusted to the actual inflation and in which output and inflation are constant. This situation is called the “steady- state” because nothing is changing. In this state the growth rate of money on the aggregate demand side, determines the inflation rate and output on aggregate supply

side is at its potential level. Steady state relationships are useful in indicating the long run behavior of the economy (Mishkin, 2022).

Changes in real income affect the exchange rate. An effect predicted by the combination of the quantity theory (quantity theory states that in any country, the money supply is equated with the demand for money which is directly proportional to the value of GNP) and the purchasing power parity.

Suppose a country's real income shifts up to a growth path of 10 percent above the path it would otherwise have followed. This event may be due to a certain discovery of wide oil reserves in their country. The extra sale of that country's oil would call forth a new demand for holding their domestic currency as a currency to facilitate purchase of that country's oil. If the extra oil results in a 10 percent rise in its national income, the quantity theory depicts a 10 percent higher transaction demand for their currency. The 10 percent income rise should push the demand out for let's say British pounds of the stock of dollars. But this extra demand cannot be met since Britain's money stock is still just as large originally as its stock of dollars. The result, the general clamor to get pounds by borrowing them or by selling Britain goods to them will lead to a rise in the values of the pound. However, income is not an independent force that can just move by itself, so caution is still necessary.

In this example, real income was being raised for a supply -side reason- Britain's ability to supply oil. It is thus easy to believe that this would strengthen the pound, either by using the quantity theory of equation or by thinking about the extra oil exports as something the other countries would need pounds to pay for (Lindert, 1991).

We see through this how income might affect exchange rate. This would give us a view on how exchange rate on the other hand might affect income or output.

4. Model

The model that the paper will be estimating is a modified equation based on the article written by Nunnenkamp et. al. (1990). A detailed discussion of the modified equation will be presented in order to facilitate the modification of the model to be estimated by the paper. After which, the variables to undergo estimation in the modified equation will be defined and explained.

The regression equation that was estimated in the paper's base article is of the form:

$$Y = a + bR + dUOG + eTOT + fDY(1)$$

Where:

Y = annual real growth of GDP per capita

R = annual changes in real effective exchange rate

$EGDP$ = percentage share of government expenditure in GDP

UOG = unanticipated annual change in money supply

TOT = terms of trade shocks

DY = three-period moving average of the once, twice and thrice lagged values of the dependent variable Y

The terms of trade is given by the following equation:

$$TOT_t = \{ [(Pm_t - Pm_{(t-1)}) MV_t] - [(Px_t - Px_{(t-1)}) XV_t] \} / (X_t + M_t) \quad (2)$$

Where:

Pm = import prices by unit value

Px = export prices by unit value

MV = imports at current prices

XV = exports at current prices

The variables $EGDP$ and UOG are introduced for economic policy measures, the growth of external shocks is accounted for by the variable TOT and pre-devaluation differences in economic growth represented by the variable DY . In addition to the growth impact of policy adjustments, the model accounts for the growth effects of factors that are beyond the control of the base articles' sample countries. Equation (1) is estimated for 48 developing countries for which pooled time-series cross-country regressions are run for 1982-1987 period using ordinary least square techniques. Regressions are also run for various sub-groups of the overall sample. In addition to equation (1), two other equations are estimated to account for lag values of the explanatory variables.

From equation (1) above, the paper has modified the equation wherein we changed the dependent variable and reduced the number of explanatory variables. The reduced form for equation (1) for which we modified for our paper is of the form:

$$RGDP_t = \alpha_1 + \beta_1 PDR_t + \beta_2 BOP_t + e_t \quad (3)$$

Where:

$RGDP_t$ = Quarterly Real GDP

PDR_t = Quarterly end period of nominal foreign exchange rate (US dollar-Ph peso exchange rate)

BOP_t = Quarterly values for BOP

e_t = stochastic error term

The model is estimated using ordinary least squares techniques for the period 2005-2025. The equation would like to estimate the impact of depreciation in the foreign exchange rate to Real GDP. The variable BOP is an additional variable to account for external shocks in the inflow and outflow of foreign currency in the Philippine economy.

5. Data and Methodology

The impact of the Peso-Dollar exchange rate and balance of payments (BOP) on real gross domestic product (RGDP) is examined using the following methodology:

$$RGDP_t = \alpha_1 + \beta_1 PDR_t + \beta_2 BOP_t + e_t \quad (4)$$

where $RGDP_t$ is the real gross domestic product at quarter t , PDR_t is the Peso-Dollar exchange rate at quarter t , BOP_t is the balance of payments at quarter t , and e_t is the error term.

The variables PDR_t and BOP_t are regressed against the dependent variable $RGDP_t$ using a vector error correction model.

The data used in this paper is on a quarterly basis. The data for the for the various variables used in this paper span a 21-year period which begins on the first quarter of 2005 and ends on the fourth and final quarter of 2025. This creates a total of 84 observations. Data on real gross domestic product (GDP), on balance of payments and the Peso-Dollar exchange rate was sourced from the *Bangko Sentral ng Pilipinas* (BSP).

Quarterly data on real gross domestic product is in millions of dollars at constant prices using 2018 as the base year (2018=100). The data on the Peso-Dollar exchange rates are the averages for the months of March, June, September and December which represent the end of the quarter from March of 2005 to December of 2025. Data on balance of payments are in millions of dollars.

6. Results

Nelson and Plosser (1982) show that the macroeconomic time series data often exhibit a unit root process and are not stationary. Results that are based on nonstationary data produce spurious results. Therefore, time series data should be tested for stationarity. I test for the stationarity of my variables by using the augmented Dickey-Fuller test. The equation for this test is

$$\Delta X_t = \alpha + \rho_t + \beta X_{t-1} + \sum_{i=1}^n \lambda_i \Delta X_{t-1} + \varepsilon_t (5)$$

In this equation, X is the variable under consideration, Δ is the first difference operator, t is a time trend, and ε is a stationary random error term. Table 1 gives us the results of the augmented Dickey Fuller test for unit roots. The results show that *RGDP* and *PDR* are not stationary at level form while *BOP* is stationary at level form. All variables are stationary in first difference form.

Table 1. Unit Root Test

Variable	Level	First Difference
<i>RGDP</i>	0.0064	-1.3729**
<i>PDR</i>	-0.0321	-0.9605***
<i>BOP</i>	-0.7333***	-7.1801***

**Significant at the 5% level

***Significant at the 1% level

Since the variables are found to be first order integrated, Johansen's (1988) cointegration test is conducted to test the long run relationship among the variables in the model. This step is conducted before estimating the model in the first difference form. This test result is presented in Table 2. The p-values in Table 2 suggest that the null hypothesis of no cointegration is rejected.

Table 2. Johansen's Cointegration Test

H_0	Eigenvalue	Trace test	p-value
$r = 0$	0.35531	50.411	0.0000
$r \leq 1$	0.16753	14.854	0.0610
$r \leq 2$	0.00003	0.002	0.9631

The following regression results were obtained using Gretl statistical software. Table 3 presents the estimates of equation (4). Before performing the regression, I hypothesized that both Peso-Dollar exchange rate and the balance of payments would affect the real gross domestic product. The formula we obtained from our model got mixed results. The sign of

ΔPDR is positive and significant at the 5% level in the short run (lags of 1 and 2 periods) which implies a positive relationship between the real gross domestic product (GDP) and the Peso-Dollar exchange rate. Therefore, there is a strong relationship between real gross domestic product and the Peso-Dollar exchange rate. ΔBOP shows a positive coefficient at t-1 and a negative coefficient at t-2, which points towards an inverse relationship with the real gross domestic product. However, the balance of payments is not statistically significant any level. Therefore, the balance of payments has minimal or negligible effect on real gross domestic product. The adjusted R-squared has a value of 0.6913 which suggests a pretty good fit for the regression equation. Therefore, the results differed from the original estimate. It turns out that only the Peso-Dollar exchange rate has a direct effect on real gross domestic product.

Table 3. *Vector Error Correction Model Estimation Results (Dependent variable: $\Delta RGDP$)*

Variables	Coefficients standard errors in ()
Constant	-1838.42 (3901.28)
$\Delta RGDP_{t-1}$	-0.7745*** (0.1170)
$\Delta RGDP_{t-2}$	-0.0958 (0.1160)
ΔBOP_{t-1}	0.0914 (0.2854)
ΔBOP_{t-2}	-0.0889 (0.2224)
ΔPDR_{t-1}	711.320** (353.756)
ΔPDR_{t-2}	753.407** (363.656)
Adjusted R ²	0.6913
Durbin-Watson	1.7553

*Significant at the 10% level

**Significant at the 5% level

***Significant at the 1% level

7. Conclusion

This paper presents the trend in economic performance represented by the quarterly GDP of the Philippines and the movements in the foreign exchange rate represented by the US dollar-Philippine peso exchange rate. The literature review of the paper summarized the movements in the US dollar-Philippine peso exchange rate and accounted for the exchange rate adjustments implemented by the government to devalue the peso to spur economic growth during the mid 1980 crisis. The paper presented how the original equation used by Nunnenkamp et. al. (1990) was modified to simplify the analysis for the case when the Philippines adopted a flexible exchange rate regime during 1992.

Based on the regression results, there are several things we can conclude about the estimated relationship between Real GDP, the foreign exchange rate and the BOP: (1) there is a positive relationship between Real GDP and the US dollar-Philippine peso exchange rate, and (2) the relationship between Real GDP and the BOP is not statistically significant at least in the Philippine context. This means that a depreciation of the domestic currency stimulates

economic activity. However, these conclusions from the estimate equation are net effects of increases in one of these variables holding other variable constant.

The paper suggests that a more accurate model could be built that captures better the relationship between economic growth and movements in the foreign exchange rate. For policy-makers, the paper suggests that a critical study of the lag response of the Philippine economy is needed so that policies to stimulate economic activity can be properly implemented.

References

- Abendin, S., and Duan, P. (2021). International Trade and Economic Growth in Africa: The Role of the Digital Economy. *Cogent Economics & Finance* 9(1): 1911767.
- Antonio, E., and Laig, M.L. (1995). The Fallacy of Exchange Rates and Exports. *Recent Economic Indicators*. 7-10.
- Conrad, D., and Jagessar, J. (2018). Real Exchange Rate Misalignment and Economic Growth: The Case of Trinidad and Tobago. *Economies* 6(4): 52.
- Cororaton, C. (1997). *Exchange Rate Movements in the Philippines* (Discussion Paper Series No. 97-05). Philippine Institute for Development Studies.
- Dornbush, R., Fisher, S. and Startz, R. (2018). *Macroeconomics* (13th ed.). McGraw Hill.
- Jawaid, S.T., and Waheed, A. (2011). Effects of Terms of Trade and its Volatility on Economic Growth: A Cross Country Empirical Investigation. *Transition Studies Review* 18(2): 217-229.
- Johansen, S. (1988). Statistical Analysis of Cointegration Vector. *Journal of Economic Dynamics and Control* 12(2): 231-254.
- Keho, Y. (2017). The Impact of Trade Openness on Economic Growth: The Case of Cote d'Ivoire. *Cogent Economics & Finance* 5(1): 1332820.
- Lindert, P.H. (1991). *International Economics* (9th ed.). McGraw-Hill.
- Miranda, C. (1995). *Determining the Foreign Exchange Rate in Aid of Export Industries* (Discussion Paper No. 9510). University of the Philippines School of Economics.
- Mishkin, F S. (2022). *The Economics of Money, Banking, and Financial Markets* (13th ed.). Pearson.
- Nelson, C.R., and Plosser, C.I. (1982). Trends and Random Walks in Macroeconomic Time Series: Some Evidence and Implications. *Journal of Monetary Economics* 10(2): 139-162
- Nunnenkamp, P., and Schweickert, R. (1990). *Real Exchange Rates and Economic Growth in Developing Countries: Is Devaluation Contractionary?* (Kiel Working Paper No. 405). Kiel Institute of World Economics.
- Parikh, A. (2006). Relationship Between Trade Liberalization, Growth, and Balance of Payments in Developing Countries: An Econometric Study. *The International Trade Journal* 20(4): 429-467.
- Pradumna, R. (1983). *The Impact of the Current Exchange Rate System on Trade and Inflation of Selected Developing Member Countries* (Economic Staff Paper No. 18). Asian Development Bank.
- Shahzad, U., He, X., and Muhammad, S. (2023). How Terms of Trade Impact Economic Growth: The Case of the United States. *Sage Open*, 13(2).
- Yap, J. (1996). *Inflation and Economic Growth in the Philippines*. (Discussion Paper No. 96-11). Philippine Institute for Development Studies.

Social sector expenditure, economic growth, and multidimensional poverty in India: evidence from panel data analysis

Balaram ROULA

Pondicherry University, India
roulabalaram1993@pondiuni.ac.in

Dr. N. KUBENDRAN

Pondicherry University, India
Kubendran.eco@pondiuni.ac.in

Abstract. *This study examines the impact of education expenditure, health expenditure, and Gross State Domestic Product (GSDP) on the Multidimensional Poverty Index (MPI) across selected Indian states during 2000–01 to 2022–23. The empirical analysis employs panel data techniques including Fixed Effect, Random Effect, Hausman specification test, and Driscoll–Kraay Fixed Effect estimator using an unbalanced panel dataset of 322 observations. The findings reveal that health expenditure has a negative and statistically significant effect on multidimensional poverty, indicating that higher public healthcare expenditure reduces deprivation across states. In contrast, GSDP shows a positive and significant relationship with MPI, suggesting that economic growth alone may not reduce poverty without inclusive development policies. However, education expenditure was found to be statistically insignificant during the study period. The study emphasizes the importance of strengthening social sector expenditure, particularly healthcare investment, for achieving inclusive and sustainable development across Indian states.*

Keywords: Multidimensional Poverty Index, Health Expenditure, Education Expenditure, GSDP, Panel Data, Driscoll–Kraay Estimator, Indian States.

JEL Classification: I32, H51, H52, O47.

1. Introduction

Multidimensional poverty has emerged as one of the major developmental challenges in both developing and underdeveloped economies. Unlike income poverty, multidimensional poverty reflects deprivation in various aspects of human life such as education, health, living standards, sanitation, housing, and access to basic services. The Multidimensional Poverty Index (MPI), developed by the United Nations Development Programme (UNDP) and Oxford Poverty and Human Development Initiative (OPHI), has become an important indicator for measuring the intensity and extent of deprivation among households. In recent years, policymakers and researchers have increasingly emphasized that economic growth alone is insufficient to reduce poverty unless supported by effective social sector expenditure and inclusive development policies.

India has experienced significant economic growth during the post-reform period, accompanied by substantial increases in Gross State Domestic Product (GSDP). However, despite rapid economic expansion, multidimensional poverty continues to persist across several Indian states due to unequal access to healthcare, education, sanitation, nutrition, and social welfare services. Considerable interstate disparities remain visible in poverty reduction and human development outcomes. States with better social infrastructure and welfare expenditure have generally performed better in reducing poverty compared to economically backward states.

Public expenditure on the social sector, particularly education and health expenditure, plays a crucial role in improving welfare conditions and reducing multidimensional deprivation. According to Wagner's theory of public expenditure, government expenditure tends to increase with economic development due to rising welfare and social needs. Similarly, endogenous growth theories highlight the importance of investment in human capital through education and healthcare for achieving long-term socio-economic development. Expenditure on health improves productivity, life expectancy, and quality of life, while education expenditure contributes to literacy, skill development, and employment generation.

Several empirical studies have examined the relationship between social sector expenditure and poverty reduction in India. Goswami and Bezbaruah (2011) found that social sector expenditure positively contributes to human development outcomes across Indian states. Agarwal (2015) observed that public expenditure on education and health significantly improves welfare indicators and reduces regional disparities. Reddy and Reddy (2019) concluded that expenditure on health, education, and welfare programmes contributes toward poverty reduction and human development improvement in India. Similarly, Das, Mandal, and Patra (2021) identified a long-run relationship between social sector expenditure and developmental outcomes across Indian states.

Although numerous studies have examined the relationship between social sector expenditure and human development, limited studies have specifically analysed the impact of education expenditure, health expenditure, and economic growth on multidimensional poverty across Indian states using robust panel data estimation techniques. Further, many earlier studies did not adequately address econometric issues such as heteroskedasticity, autocorrelation, and cross-sectional dependence in panel datasets.

In this context, the present study attempts to examine the impact of education expenditure, health expenditure, and Gross State Domestic Product on Multidimensional Poverty Index across selected Indian states during the period 2000–01 to 2022–23. The study employs panel data estimation techniques including Fixed Effect, Random Effect, Hausman specification test, and Driscoll–Kraay Fixed Effect estimator to obtain robust and reliable empirical estimates. The findings of the study are expected to provide important policy implications regarding the role of social sector expenditure in reducing multidimensional poverty and promoting inclusive development across Indian states.

2. Review of Literature

A large number of empirical studies have examined the relationship between social sector expenditure, economic growth, and poverty reduction in both developed and developing economies. In the Indian context, social sector expenditure has increasingly been recognised as an important instrument for improving welfare conditions, reducing deprivation, and promoting inclusive development. Public expenditure on education, health, sanitation, nutrition, and welfare programmes plays a significant role in enhancing human capabilities and reducing multidimensional poverty across states.

Goswami and Bezbaruah (2011) analysed the impact of social sector expenditure on Human Development Index (HDI) across Indian states during the post-reform period. The study found that although social expenditure increased over time, interstate differences in human development were influenced more by per capita income than by expenditure itself. Agarwal (2015) examined the relationship between social sector expenditure and human development in India and observed that education and health expenditure positively influenced development outcomes, though interstate disparities remained persistent. Similarly, Mittal (2016) argued that higher per capita social sector expenditure significantly improved HDI across Indian states, particularly in states such as Kerala and Tamil Nadu which consistently invested more in health and education.

Reddy and Reddy (2019) employed Tobit regression analysis to study the impact of social sector expenditure on human development in India. The findings revealed that expenditure on education, healthcare, sanitation, housing, nutrition, and welfare programmes positively contributed toward improving human development indicators. Likewise, Dutt, Das, and De (2023) observed that increased social sector expenditure improved education, healthcare, and living standards, thereby contributing to both human development and economic growth.

Several studies have specifically focused on the role of education and health expenditure in poverty reduction and welfare improvement. Ruzima and Veerachamy (2023) analysed the long-run relationship between public expenditure on health, education, and human development in India during 1990–2018. The study found that health expenditure had a positive and significant impact on human development, whereas the effect of education expenditure was mixed due to issues such as poor educational quality and high dropout rates. Sinha (2025) also highlighted that expenditure on education and health contributes significantly toward human capital formation, productivity growth, and inclusive economic development.

In the context of poverty and deprivation, Alkire and Foster (2011) developed the multidimensional poverty measurement framework, which has been widely adopted by the United Nations Development Programme (UNDP) for estimating the Multidimensional Poverty Index (MPI). Their approach emphasised that poverty extends beyond income deprivation and includes deficiencies in health, education, and living standards. The MPI approach has increasingly gained policy relevance in developing countries including India.

Several Indian studies have examined multidimensional poverty and regional disparities. Seth and Alkire (2014) analysed multidimensional poverty across Indian states and observed substantial interstate disparities in deprivation levels. The study highlighted that states with better social infrastructure and higher welfare expenditure performed relatively better in reducing multidimensional poverty. Mohanty and Vasishtha (2018) also found that multidimensional poverty in India is closely associated with inadequate access to healthcare, education, sanitation, and housing facilities.

Ranjan and Panda (2021) studied development expenditure and its relationship with HDI and GSDP in low-income states of India. The study observed that although development expenditure contributed positively toward economic growth, its long-run impact on human development remained limited due to unequal distribution of resources. Das, Mandal, and Patra (2021) employed panel data analysis to examine the relationship between social sector expenditure and human development across Indian states and found the existence of long-run associations between government expenditure and development indicators.

A number of studies have also investigated regional disparities and efficiency of social sector expenditure across Indian states. Chattopadhyay (2017) argued that despite policy emphasis on inclusive growth, social sector expenditure remained inadequate in several Indian states during the post-reform period. Soni (2019) identified considerable interstate variations in health and education outcomes, with states such as Kerala and Himachal Pradesh performing better due to higher social sector investment. Purohit (2014) assessed the efficiency of social sector expenditure using Free Disposable Hull (FDH) and Stochastic Frontier Analysis (SFA) techniques and found significant disparities in the utilization efficiency of health and education expenditure across states.

Roula and Kumaran (2025) analysed public expenditure efficiency across Indian states using Data Envelopment Analysis (DEA). The study reported that economically advanced states with better governance and infrastructure demonstrated higher efficiency in public expenditure utilization compared to poorer states. Similarly, Bose and Banerjee (2025) examined the relationship between social spending, fiscal policy, and human development in India and argued that fiscal consolidation policies under the FRBM framework constrained public expenditure on education and health despite rising economic growth.

Further, Roula (2025) investigated the role of social sector expenditure in rural empowerment and found that expenditure on health, education, nutrition, and welfare programmes significantly reduced rural unemployment and improved wage levels. Roula, Sethi, and Sahu (2026) also emphasised the importance of fiscal discipline and efficient public expenditure management for achieving sustainable development outcomes across states.

Although extensive literature exists on social sector expenditure, human development, and economic growth, relatively few studies have specifically examined the impact of education expenditure, health expenditure, and economic growth on Multidimensional Poverty Index (MPI) across Indian states using robust panel data techniques. Moreover, earlier studies have often ignored econometric issues such as heteroskedasticity, serial correlation, and cross-sectional dependence in macro panel datasets. Therefore, the present study attempts to fill this research gap by employing robust panel data estimation methods such as the Driscoll–Kraay Fixed Effect estimator to analyse the impact of education expenditure, health expenditure, and GSDP on multidimensional poverty across selected Indian states.

The existing literature largely focuses on the relationship between social sector expenditure, human development, economic growth, and poverty reduction in India. Several studies have examined the impact of education and health expenditure on Human Development Index (HDI), welfare indicators, and economic growth across Indian states. Other studies have analysed regional disparities, efficiency of public expenditure, and the role of fiscal policy in promoting inclusive development. However, relatively limited empirical research has specifically investigated the impact of education expenditure, health expenditure, and Gross State Domestic Product (GSDP) on Multidimensional Poverty Index (MPI) across Indian states.

Further, most earlier studies have concentrated either on income poverty, human development outcomes, or overall development expenditure without adequately addressing multidimensional poverty from a panel data perspective. Moreover, many previous studies have not employed robust econometric techniques capable of addressing important panel data issues such as heteroskedasticity, autocorrelation, and cross-sectional dependence, which may affect the reliability of empirical estimates.

In this context, the present study attempts to bridge this research gap by examining the impact of education expenditure, health expenditure, and economic growth on multidimensional poverty across selected Indian states during 2000–01 to 2022–23. The study employs robust panel data estimation techniques including Fixed Effect, Random Effect, Hausman specification test, and Driscoll–Kraay Fixed Effect estimator to obtain reliable and consistent empirical results. Therefore, the study contributes to the existing literature by providing a comprehensive panel data analysis of the determinants of multidimensional poverty in India.

3. Data Source

The present study is based entirely on secondary data collected from various official government publications and databases. The analysis covers 14 major Indian states for the period from 2000–01 to 2022–23, forming an unbalanced panel dataset consisting of 322 observations.

Data relating to the Multidimensional Poverty Index (MPI) have been obtained from reports published by NITI Aayog, the United Nations Development Programme (UNDP),

and various State Human Development Reports. Information on education expenditure and health expenditure has been collected from the Reserve Bank of India (RBI) publication *State Finances: A Study of Budgets* along with respective state budget documents. Data on Gross State Domestic Product (GSDP) have been sourced from the Ministry of Statistics and Programme Implementation (MOSPI), Government of India.

Table 1: *Variables, Measurement, and Data Sources*

Variables	Description	Measurement/Proxy	Source
MPI	Multidimensional Poverty Index	Index value measuring deprivation in health, education, and living standards	NITI Aayog, UNDP, State Human Development Reports
EE	Education Expenditure	Total government expenditure on education	RBI, <i>State Finances: A Study of Budgets</i> and State Budget Documents
HE	Health Expenditure	Total government expenditure on health sector	RBI, <i>State Finances: A Study of Budgets</i> and State Budget Documents
GSDP	Gross State Domestic Product	Gross State Domestic Product at current prices	Ministry of Statistics and Programme Implementation (MOSPI)
LNEE	Logarithm of Education Expenditure	ln(EE)	Computed by Author
LNHE	Logarithm of Health Expenditure	ln(HE)	Computed by Author
LNGSDP	Logarithm of Gross State Domestic Product	ln(GSDP)	Computed by Author

Source: Compiled by the author based on various official publications.

The study uses education expenditure, health expenditure, and GSDP as the major explanatory variables influencing multidimensional poverty across states. To reduce scale differences, improve normality, minimize heteroskedasticity, and facilitate elasticity interpretation of coefficients, these variables have been transformed into logarithmic form before empirical estimation.

The study further employs panel econometric techniques, including Fixed Effect and Random Effect models, to analyse the relationship between social sector expenditure and multidimensional poverty across Indian states. In addition, several diagnostic and robustness tests have been conducted to ensure the reliability and consistency of the estimated results.

4. Model Specification

The present study examines the impact of social sector expenditure on multidimensional poverty across Indian states using panel data regression analysis. Panel data methodology is considered appropriate because it combines both cross-sectional and time-series dimensions, thereby enabling the study to capture state-specific heterogeneity and temporal variations more effectively than pure cross-sectional or time-series techniques. The study constructs an unbalanced panel dataset comprising 14 major Indian states covering the period from 2000–01 to 2022–23, with a total of 322 observations.

The functional relationship between Multidimensional Poverty Index (MPI) and the explanatory variables is expressed as:

$$MPI = f(LNEE, LNHE, LNGSDP)$$

The econometric model of the study is specified as:

$$MPI_{it} = \beta_0 + \beta_1 LNEE_{it} + \beta_2 LNHE_{it} + \beta_3 LNGSDP_{it} + \mu_i + \varepsilon_{it}$$

where MPI_{it} denotes the Multidimensional Poverty Index of state i at time t . The variable $LNEE_{it}$ represents the logarithm of education expenditure, while $LNHE_{it}$ denotes the logarithm of health expenditure. Further, $LNGSDP_{it}$ indicates the logarithm of Gross State Domestic Product. In the model, β_0 represents the intercept term, whereas β_1 , β_2 , and β_3 are the slope coefficients of the explanatory variables. The term μ_i captures unobserved state-specific effects, and ε_{it} represents the random error term.

The explanatory variables are transformed into logarithmic form to reduce scale variations, improve normality, minimize heteroskedasticity, and facilitate elasticity interpretation of coefficients. The logarithmic transformation equations are expressed as:

$$LNEE = \ln(EE)$$

$$LNHE = \ln(HE)$$

$$LNGSDP = \ln(GSDP)$$

where EE denotes education expenditure, HE represents health expenditure, and $GSDP$ refers to Gross State Domestic Product.

Initially, both Fixed Effect (FE) and Random Effect (RE) panel regression models are estimated. The Fixed Effect model assumes that state-specific effects are correlated with the explanatory variables and controls for unobserved time-invariant heterogeneity across states. The Fixed Effect model is represented as:

$$MPI_{it} = \alpha_i + \beta_1 LNEE_{it} + \beta_2 LNHE_{it} + \beta_3 LNGSDP_{it} + \varepsilon_{it}$$

where α_i represents state-specific intercepts.

In contrast, the Random Effect model assumes that individual state-specific effects are random and uncorrelated with the explanatory variables. The Random Effect specification is expressed as:

$$MPI_{it} = \beta_0 + \beta_1 LNEE_{it} + \beta_2 LNHE_{it} + \beta_3 LNGSDP_{it} + u_i + \varepsilon_{it}$$

To determine the appropriate estimation technique between the Fixed Effect and Random Effect models, the Hausman Specification Test is employed. The null hypothesis of the Hausman test states that the Random Effect model is appropriate, while the alternative hypothesis supports the Fixed Effect model. The Hausman test statistic is given as:

$$H = (\hat{\beta}_{FE} - \hat{\beta}_{RE})' [Var(\hat{\beta}_{FE}) - Var(\hat{\beta}_{RE})]^{-1} (\hat{\beta}_{FE} - \hat{\beta}_{RE})$$

Further, the Levin–Lin–Chu (LLC) panel unit root test is applied to examine the stationarity properties of the variables. The general LLC test equation is expressed as:

$$\Delta Y_{it} = \alpha Y_{it-1} + \sum_{j=1}^p \beta_j \Delta Y_{it-j} + \mu_i + \varepsilon_{it}$$

The study also employs the Kao panel cointegration test to investigate the existence of long-run equilibrium relationships among the variables.

In addition, diagnostic tests are conducted to examine the presence of heteroskedasticity, autocorrelation, and cross-sectional dependence. The Modified Wald test is used to detect heteroskedasticity, the Wooldridge test is employed for autocorrelation, and the Pesaran Cross-Sectional Dependence (CD) test is applied to identify cross-sectional dependence across states.

Since the empirical results indicate the presence of heteroskedasticity, autocorrelation, and cross-sectional dependence, the study finally adopts the Driscoll–Kraay Fixed Effect estimator for robust estimation. The Driscoll–Kraay estimator provides standard errors that are robust to heteroskedasticity, serial correlation, and cross-sectional dependence simultaneously, making it particularly suitable for macro panel datasets involving Indian states.

5. Results and Discussion

Table 2: *Descriptive Statistics of Variables*

Variables	Observations	Mean	Standard Deviation	Minimum	Maximum
MPI	322	0.1875	0.1400	0.0020	0.5590
LNEE	322	14.0043	0.6787	12.5708	15.7604
LNHE	322	12.9311	0.8728	10.9194	15.0280
LNGSDP	322	17.4954	0.7233	15.8813	19.2277

Sources: Author's own estimation using Stata.

The descriptive statistics indicate moderate variation among the variables across selected Indian states during the study period. The average value of Multidimensional Poverty Index (MPI) is 0.1875, with values ranging from 0.0020 to 0.5590, indicating substantial interstate differences in multidimensional poverty. The mean values of logarithm of education expenditure (LNEE), logarithm of health expenditure (LNHE), and logarithm of Gross State Domestic Product (LNGSDP) are 14.0043, 12.9311, and 17.4954 respectively. The standard deviations suggest reasonable dispersion in public expenditure and economic growth across states.

Table 3: *Correlation Matrix of Variables*

Variables	MPI	LNEE	LNHE	LNGSDP
MPI	1.0000			
LNEE	-0.4689***	1.0000		
LNHE	-0.6192***	0.9356***	1.0000	
LNGSDP	-0.5381***	0.9658***	0.9909***	1.0000

Note: *** indicates significance at 1 percent level.

Sources: Author's own estimation using Stata.

The correlation matrix indicates that Multidimensional Poverty Index (MPI) is negatively associated with education expenditure (LNEE), health expenditure (LNHE), and Gross State Domestic Product (LNGSDP). The negative coefficients imply that higher social sector expenditure and economic growth are associated with lower multidimensional poverty across Indian states. Among the explanatory variables, health expenditure exhibits the strongest negative association with MPI. Further, the explanatory variables are highly positively correlated with one another, indicating close association among fiscal and economic indicators.

Table 4: *Levin–Lin–Chu (LLC) Panel Unit Root Test*

Variables	Adjusted t-statistic	P-value	Order of Integration
MPI	-2.1168	0.0171	Stationary at Level
LNEE	-5.4126	0.0000	Stationary at Level
LNHE	-3.6146	0.0002	Stationary at Level
LNGSDP	-3.6827	0.0001	Stationary at Level

Sources: Author's own estimation using Stata.

The Levin–Lin–Chu (LLC) panel unit root test was employed to examine the stationarity properties of the variables used in the study. The null hypothesis of the LLC test states that the panels contain unit roots, while the alternative hypothesis suggests stationarity.

The results reveal that the adjusted t-statistics of MPI, LNEE, LNHE, and LNGSDP are statistically significant with p-values less than 0.05. Therefore, the null hypothesis of unit root is rejected for all variables. This indicates that all variables are stationary at level and integrated of order zero, $I(0)$. Hence, the variables do not suffer from non-stationarity problems, validating the application of panel regression techniques such as Fixed Effect and Random Effect models.

Table 5: *Kao Panel Cointegration Test*

Test Statistics	Statistic Value	P-value
Modified Dickey–Fuller t	1.5837	0.0566
Dickey–Fuller t	1.6358	0.0509
Augmented Dickey–Fuller t	1.1947	0.1161
Unadjusted Modified Dickey–Fuller t	1.5408	0.0617
Unadjusted Dickey–Fuller t	1.5866	0.0563

Sources: Author’s own estimation using Stata.

The Kao panel cointegration test was employed to examine the long-run equilibrium relationship among Multidimensional Poverty Index (MPI), education expenditure (LNEE), health expenditure (LNHE), and Gross State Domestic Product (LNGSDP). The null hypothesis of the Kao test states that there is no cointegration among the variables, while the alternative hypothesis suggests that all panels are cointegrated.

The empirical findings indicate that the Dickey–Fuller t-statistic is marginally significant at the 10 percent level with a p-value of 0.0509. Similarly, the Modified Dickey–Fuller and Unadjusted Dickey–Fuller statistics also show near-significant values. Although the Augmented Dickey–Fuller statistic is statistically insignificant, the majority of the test statistics provide evidence supporting the existence of a long-run equilibrium relationship among the variables.

Therefore, the study suggests that education expenditure, health expenditure, and economic growth maintain a long-run association with multidimensional poverty across selected Indian states. This validates the use of panel regression estimation techniques for further empirical analysis.

Table 6: *Fixed Effect and Random Effect Model Results (MPI Model)*

Variables	Fixed Effect (FE)	Random Effect (RE)
LNEE	0.0222 (0.0311)	0.0007 (0.0306)
LNHE	-0.5248*** (0.0742)	-0.6001*** (0.0642)
LNGSDP	0.4881*** (0.1133)	0.6063*** (0.1003)
_cons	-1.8773** (0.7937)	-2.6685*** (0.6726)
Observations	322	322
Number of Groups	14	14
Within R ²	0.8572	0.8567
Between R ²	0.3544	0.4978

Variables	Fixed Effect (FE)	Random Effect (RE)
Overall R ²	0.6479	0.6809
F/Wald Statistic	610.27	1799.65
Prob > Test	0.0000	0.0000
Hausman Test (p-value)	\multicolumn{2}{c}{0.0016}	

Note: *** p<0.01, ** p<0.05, * p<0.10.

Sources: Author's own estimation using Stata.

The table presents the Fixed Effect (FE) and Random Effect (RE) estimation results for the MPI model across selected Indian states. The Hausman specification test reports a p-value of 0.0016, which is statistically significant at the 1 percent level. Therefore, the null hypothesis of the Random Effect model is rejected, and the Fixed Effect model is considered appropriate for empirical analysis.

The FE estimation results reveal that health expenditure (LNHE) has a negative and statistically significant effect on Multidimensional Poverty Index (MPI). The coefficient value of -0.5248 implies that an increase in health expenditure significantly reduces multidimensional poverty across states. This finding suggests that greater public investment in healthcare improves living conditions, health accessibility, and overall welfare, thereby reducing deprivation among households.

On the other hand, Gross State Domestic Product (LNGSDP) exhibits a positive and statistically significant relationship with MPI. The coefficient value of 0.4881 indicates that economic growth alone may not necessarily reduce multidimensional poverty unless accompanied by inclusive development policies and equitable distribution of income. The result suggests that higher economic growth in some states may coexist with persistent poverty and inequality.

Further, education expenditure (LNEE) shows a positive but statistically insignificant relationship with MPI. This indicates that education expenditure alone may not have an immediate measurable impact on multidimensional poverty reduction during the study period.

The model exhibits strong explanatory power with a within R² value of 0.8572, implying that approximately 85.72 percent of the variation in MPI within states is explained by the explanatory variables included in the model.

Table 7: Diagnostic Test Results for MPI Model

Diagnostic Test	Null Hypothesis	Test Statistic	P-value	Decision
Hausman Test	Random Effect model is appropriate	$\chi^2 = 15.31$	0.0016	Reject H ₀ ; Fixed Effect model preferred
Modified Wald Test for Heteroskedasticity	Homoskedasticity exists	$\chi^2(14) = 253.16$	0.0000	Reject H ₀ ; Heteroskedasticity present
Wooldridge Test for Autocorrelation	No first-order autocorrelation	$F(1,13) = 516.870$	0.0000	Reject H ₀ ; Autocorrelation present
Pesaran Cross-Sectional Dependence Test	No cross-sectional dependence	CD = 13.083	0.0000	Reject H ₀ ; Cross-sectional dependence present

Sources: Author's own estimation using Stata.

The diagnostic test results reveal the presence of several econometric issues in the panel dataset. The Hausman specification test confirms that the Fixed Effect model is more appropriate than the Random Effect model. Further, the Modified Wald test indicates the

existence of heteroskedasticity across panels. The Wooldridge test confirms the presence of first-order serial correlation, while Pesaran's test reveals significant cross-sectional dependence among states.

Since the panel dataset suffers from heteroskedasticity, autocorrelation, and cross-sectional dependence simultaneously, the conventional Fixed Effect estimator may produce biased and inefficient standard errors. Therefore, the study employs the Driscoll–Kraay Fixed Effect estimator to obtain robust and reliable empirical estimates.

Table 8: *Driscoll–Kraay Fixed Effect Regression Results (MPI Model)*

Variables	Coefficient	Driscoll–Kraay Std. Error	t-value	P-value
LNEE	0.0222	0.0570	0.39	0.700
LNHE	-0.5248***	0.0654	-8.03	0.000
LNGSDP	0.4881***	0.1131	4.32	0.000
_cons	-1.8773**	0.7767	-2.42	0.024

Model Statistics

Statistics	Value
Number of Observations	322
Number of Groups	14
F-statistic	206.73
Prob > F	0.0000
Within R ²	0.8572
Maximum Lag	2

Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Sources: Author's own estimation using Stata.

To obtain robust and reliable estimates in the presence of heteroskedasticity, autocorrelation, and cross-sectional dependence, the study employed the Driscoll–Kraay Fixed Effect estimator. The estimator provides robust standard errors that are consistent in the presence of these econometric issues.

The empirical findings reveal that health expenditure (LNHE) has a negative and statistically significant impact on the Multidimensional Poverty Index (MPI). The coefficient value of -0.5248 indicates that a one percent increase in health expenditure significantly reduces multidimensional poverty across Indian states. The result suggests that increased public investment in healthcare improves health accessibility, living standards, and overall welfare conditions, thereby reducing deprivation among households.

On the other hand, Gross State Domestic Product (LNGSDP) exhibits a positive and statistically significant relationship with MPI. The coefficient value of 0.4881 implies that economic growth alone does not automatically reduce multidimensional poverty unless supported by inclusive development policies and equitable distribution of income. In some states, higher economic growth may coexist with persistent inequality and deprivation.

Further, education expenditure (LNEE) shows a positive but statistically insignificant relationship with MPI. This suggests that education expenditure alone may not have an immediate measurable effect on multidimensional poverty reduction during the study period.

The model is statistically significant overall, as indicated by the F-statistic value of 206.73 with a probability value of 0.0000. Additionally, the within R^2 value of 0.8572 implies that

approximately 85.72 percent of the variation in MPI within states is explained by the explanatory variables included in the model. Therefore, the Driscoll–Kraay estimator provides robust and reliable estimates for examining the determinants of multidimensional poverty across Indian states.

6. Summary of Findings

The present study examined the impact of education expenditure, health expenditure, and Gross State Domestic Product (GSDP) on the Multidimensional Poverty Index (MPI) across selected Indian states during the period 2000–01 to 2022–23 using panel data analysis. The study employed descriptive statistics, correlation analysis, panel unit root tests, cointegration analysis, Fixed Effect and Random Effect models, diagnostic tests, and the Driscoll–Kraay Fixed Effect estimator.

The descriptive statistics revealed considerable interstate variations in multidimensional poverty, education expenditure, health expenditure, and economic growth. The correlation analysis showed that MPI is negatively associated with education expenditure, health expenditure, and GSDP, indicating that higher public expenditure and economic growth are generally linked with lower multidimensional poverty.

The Levin–Lin–Chu (LLC) panel unit root test confirmed that all variables are stationary at level and integrated of order zero, $I(0)$. Further, the Kao panel cointegration test suggested the existence of a long-run equilibrium relationship among MPI, education expenditure, health expenditure, and GSDP.

The empirical analysis estimated both Fixed Effect and Random Effect models. The Hausman specification test confirmed that the Fixed Effect model is more appropriate for the study because state-specific effects are correlated with explanatory variables. Diagnostic tests further identified the presence of heteroskedasticity, first-order autocorrelation, and cross-sectional dependence in the panel dataset. Therefore, the Driscoll–Kraay Fixed Effect estimator was employed to obtain robust and reliable estimates.

The robust regression results revealed that health expenditure has a negative and statistically significant impact on MPI, implying that higher public investment in healthcare significantly reduces multidimensional poverty across Indian states. The findings indicate that improved healthcare expenditure enhances living standards, welfare conditions, and accessibility to health services, thereby reducing deprivation among households.

In contrast, Gross State Domestic Product exhibited a positive and statistically significant relationship with MPI. The result suggests that economic growth alone may not automatically reduce multidimensional poverty unless accompanied by inclusive development strategies, equitable income distribution, and effective social welfare policies.

Further, education expenditure showed a statistically insignificant relationship with MPI, implying that education expenditure alone may not produce an immediate measurable effect on multidimensional poverty reduction during the study period.

Overall, the study concludes that social sector expenditure, particularly health expenditure, plays a crucial role in reducing multidimensional poverty across Indian states. Sustainable poverty reduction therefore requires not only economic growth but also inclusive and welfare-oriented public expenditure policies.

7. Conclusion and Policy Implications

The study examined the impact of education expenditure, health expenditure, and Gross State Domestic Product (GSDP) on Multidimensional Poverty Index (MPI) across selected Indian states during 2000–01 to 2022–23 using panel data techniques. The empirical analysis employed Fixed Effect, Random Effect, and Driscoll–Kraay Fixed Effect estimation methods.

The findings reveal that health expenditure has a negative and statistically significant impact on MPI, indicating that higher public expenditure on healthcare helps reduce multidimensional poverty across states. Improved healthcare services enhance welfare conditions, reduce deprivation, and improve living standards. On the other hand, GSDP shows a positive and significant relationship with MPI, suggesting that economic growth alone may not reduce poverty unless supported by inclusive development policies and equitable distribution of income. Further, education expenditure was found to be statistically insignificant in influencing multidimensional poverty during the study period.

The diagnostic tests confirmed the presence of heteroskedasticity, autocorrelation, and cross-sectional dependence in the panel dataset. Therefore, the Driscoll–Kraay estimator was used to obtain robust and reliable estimates.

The study suggests that policymakers should increase public expenditure on healthcare and strengthen social welfare programmes to reduce multidimensional poverty effectively. Inclusive growth policies, equitable resource distribution, improvement in educational quality, and efficient utilization of social sector expenditure are essential for achieving sustainable and balanced development across Indian states.

References

- Alkire, S., and Foster, J. (2011). Counting and multidimensional poverty measurement. *Journal of Public Economics*, 95(7–8), 476–487.
- Agarwal, P. (2015). Social sector expenditure and human development: Empirical analysis of Indian states. *Indian Journal of Human Development*, 9(2), 173–189.
- Bose, S., and Banerjee, S. (2025). Social spending and fiscal policy in India: Towards an alternate macro-fiscal framework integrating human development. *National Institute of Public Finance and Policy*.
- Chattopadhyay, S. (2017). Social sector expenditure in India in the 2000s: Trends and implications. *Journal of Development Policy and Practice*, 3(1), 16–40.

- Das, R.C., Mandal, C., and Patra, A.K. (2021). Linkage between social sector spending and HDI: Study on individual as well as panel data of Indian states. *Review of Social Economy*, 79(2), 357–379.
- Dutt, P., Das, E., and De, A. (2023). “Analysing the impact of social sector expenditure on human development and growth: A state-level study in India”. In *Aapoori: The Annual Economics Journal of Miranda House* (pp. 56–73). University of Delhi.
- Goswami, B., and Bezbaruah, M.P. (2011). Social sector expenditures and their impact on human development: The Indian experience. *Indian Journal of Human Development*, 5(2), 365–385.
- Mittal, P. (2016). Social sector expenditure and human development of Indian states. *MPRA Paper No. 75804*.
- Mohanty, S.K., and Vasishtha, G. (2018). Multidimensional poverty and deprivation in India: Regional analysis and policy implications. *Social Indicators Research*, 140(2), 421–442.
- Purohit, B.C. (2014). Efficiency of social sector expenditure in India: A case of health and education in selected Indian states. *Healthcare in Low-resource Settings*, 2(1).
- Ranjan, P., and Panda, P.K. (2021). Pattern of development spending and its impact on human development index and gross state domestic product in low-income states in India. *Journal of Development Policy and Practice*, 7(1), 71–95.
- Reddy, S.M., and Reddy, V.V.N. (2019). Social sector expenditures and their impact on human development in India: Empirical evidence during 2001–02 to 2015–16. *International Journal of Recent Technology and Engineering*, 7(6), 1938–1943.
- Roula, B. (2025). Role of social sector expenditure in empowering rural India: An analytical study. *International Journal for Multidisciplinary Research (IJFMR)*, 7(3), 1–15.
- Roula, B., and Kumaran, D. N. (2025). Evaluating the efficiency of public expenditure across Indian states. *European Economic Letters*, 15(2), 4134–4148.
- Roula, B., Sethi, S.K., and Sahu, T.K. (2026). Fiscal management and fiscal sustainability: A comparative study of Odisha and West Bengal. *ShodhKosh: Journal of Visual and Performing Arts*, 7(7s), 537–549.
- Ruzima, M., and Veerachamy, P. (2023). The impact of public spending in education and health on human development in India. *Journal of the Asia Pacific Economy*, 28(2), 390–403.
- Seth, S., and Alkire, S. (2014). Measuring and decomposing inequality among the multidimensionally poor using ordinal variables: A counting approach. *OPHI Working Paper Series*.
- Sinha, J.K. (2025). The growth effects of social sector spending in India. *Economics*, 1(1), 55–84.
- Soni, M. (2019). Social sector expenditure in India: A regional analysis. *Research Review International Journal of Multidisciplinary*, 4(6), 1–8.

Capitalization of Romanian companies and macroeconomic variables: friends or foes?*

Bogdan-Octavian COZMÂNCĂ

Bucharest University of Economic Studies, Romania
bogdan.cozmanca@fin.ase.ro

Leonard-Dan UZUM

Bucharest University of Economic Studies, Romania
uzumleonard16@stud.ase.ro

Abstract. *Adequate capitalization of non-financial companies represents one of the main vulnerabilities of the Romanian economy. Estimating its impact on macroeconomic variables could give hints about how transmission mechanisms of fiscal and monetary policies as well as financial stability could be strengthened. Using empirical evidence for Romania, in the period 2009-2024, the paper investigates the relationship between capitalization needs, economic growth and budget deficit and shows that by reducing undercapitalization, budget deficit could decrease by around 0.3-0.4 percentage points – non-ESA2010/consolidation effects, operating directly through higher taxes of undercapitalized companies and indirectly through second-round effects of more buoyant economic growth. Moreover, it is proved that adequately capitalized firms are more resilient, less dependent on external financing, and better able to sustain investment and absorb shocks. Strengthening corporate capital buffers may therefore support policy effectiveness, macroeconomic resilience, and sustainable economic growth.*

Keywords: companies' capitalization, transmission mechanism, economic growth, budget deficit.

JEL Classification: G32, O47, E62, H62, O16.

* Generative Artificial Intelligence (AI)

We confirm that in accordance with the Taylor & Francis AI Policy, we used Generative AI tools (ChatGPT) in the preparation of our manuscript to enhance the language of our manuscript, improving its clarity and readability.

1. Introduction

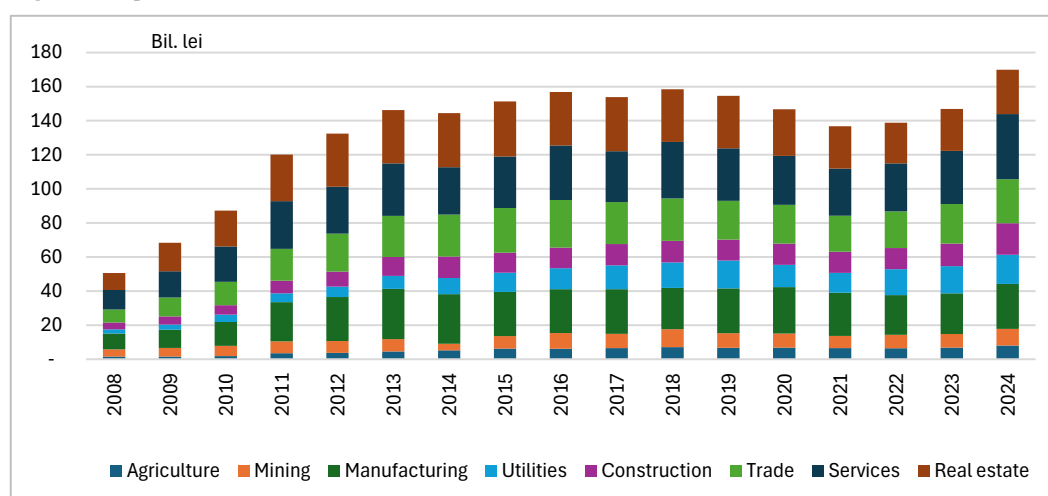
The continuation of the fiscal consolidation process, necessary to reduce the imbalances affecting the national economy, depends on increasing tax revenues. This goal can be achieved by broadening the tax base and/or improving tax collection: (i) in the case of VAT, where Romania has the largest VAT gap among EU Member States, at around 30%⁽¹⁾, significant reduction in the VAT collection gap could generate additional revenues of more than 1.5% of GDP, and (ii) in the case of corporate income tax, where the collection gap is likewise the highest in the EU, at 43.6%⁽²⁾, improved collection can generate at least 0.5% of GDP. In addition to changes in tax legislation, following the example of other countries in the region that have significantly reduced tax evasion and the shadow economy (Romania records one of the highest shares of the shadow economy in GDP among EU Member States, at around 29% of GDP⁽³⁾, after Bulgaria at 33.1% and Croatia at 29.7%), it is also necessary to strengthen the financial health of companies, in line with Recommendation R/2/2018 of the National Committee for Macprudential Oversight. Additional legislative changes, that could eliminate the significant existing arbitrage opportunity between capital and labour taxation, as well as curbing the extreme regressivity of the Romanian tax system – that has an additional benefit of higher labour market participation⁽⁴⁾ – would also partially function by improving the real sector capitalisation.

Adequate capitalization of the corporate sector is a fundamental condition for the efficient functioning of the macroeconomic policies transmission mechanisms, as the soundness of companies' financial positions and balance sheets directly influences their capacity to absorb shocks, access sources of financing and adapt decisions on investment, production, employment and financial structure in response to changes in the economic policy framework. From this perspective, an adequate level of equity capital contributes to reducing the financial vulnerability of firms and increasing their capacity to respond to the incentives generated by fiscal, monetary and macroprudential policies. Conversely, the prevalence of a high degree of undercapitalization can affect the effectiveness of these policies by limiting the capacity of companies to transform economic policy impulses into real investment and development decisions, by restricting access to credit and by amplifying the constraints associated with high levels of indebtedness. Thus, the capitalization of companies should not be analyzed exclusively as a microeconomic characteristic of the corporate financial structure, but also as a determinant of the functioning of the channels through which macroeconomic policy measures are transmitted to the real economy, with direct implications for economic growth, financial stability and the sustainability of the fiscal position. According to the June 2026 edition of the Financial Stability Report the main vulnerability of the national economy is represented by the low payment discipline in the economy and the vulnerabilities in the balance sheets of companies. Undercapitalized companies invest less, operate with fragile liquidity and are more exposed to insolvency, which reduces the tax base and increases the risk of non-payment and evasion. Moreover, the low level of capitalization affects the ability of companies to access external sources of financing, limiting both investment possibilities

and financial resilience in times of stress. Restricted access to bank lending and other market financing mechanisms contributes to the perpetuation of a financing model based predominantly on domestic funds or trade credit.

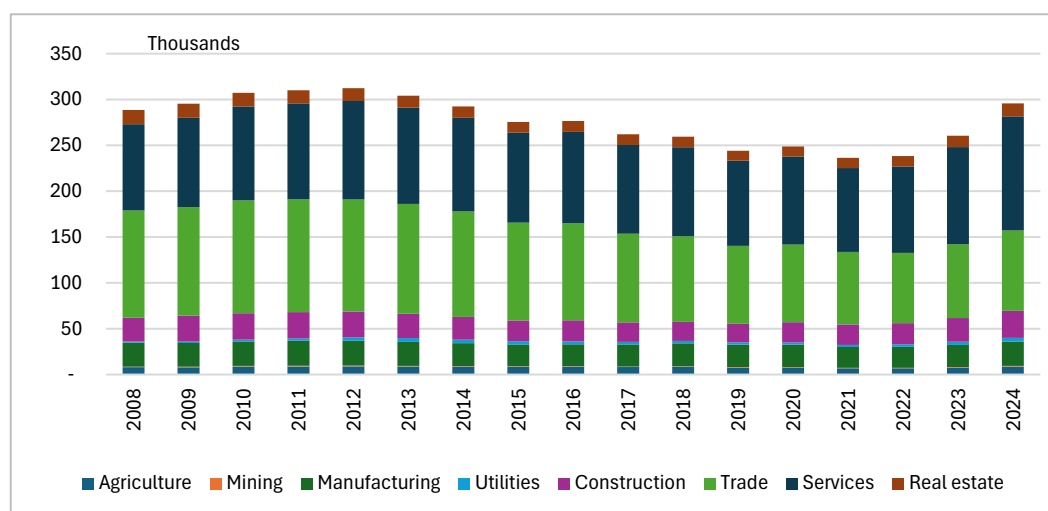
According to Romanian Ministry of Finance data, the economy-wide recapitalization need of Romanian companies amounted to RON 170 billion in 2024 (Figure 1), equivalent to approximately 9% of GDP, and was generated by nearly one-third of all companies in the economy (Figure 2).

Figure 1. Capitalization needs



Source: Ministry of Finance, authors calculations

Figure 2. Number of undercapitalized companies

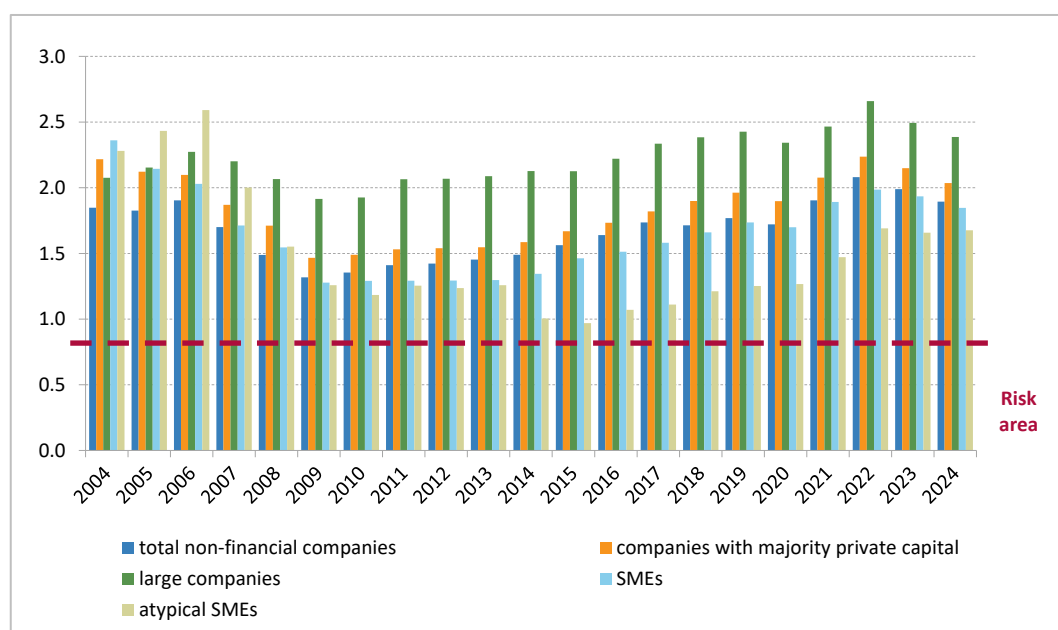


Source: Ministry of Finance, authors calculations

These companies also play a significant role in economic activity, contributing around 7% of the gross value added, holding approximately 10% of total assets, and employing more than 12% of the total number of employees. From a sectoral perspective, services account for the largest share of recapitalization needs, at 23%, followed by manufacturing at 16%, while trade and real estate account for approximately 15%.

Regarding the financial health indicator of the Romanian companies, using the Altman (2000) methodology as a proxy, the data points towards a deterioration in 2024 compared with the previous year, although it remained above the 1.2 threshold – associated with the distress zone (Figure 3). Nevertheless, several sectors continue to remain within the risk zone, including agriculture, utilities, and real estate.

Figure 3. Financial health measure by company size



Note: The overall financial health measure of the non-financial companies sector was determined based on the methodology developed by Edward Altman – “Predicting financial distress of companies: Revisiting the Z-score and ZETA models”, 2000. Considering that the non-financial companies sector in Romania comprises overwhelmingly unlisted firms, the indicator was calculated by using the method adapted for private companies: Z-score was built based on the following formula: $Z' = 0.717(X1) + 0.847(X2) + 3.107(X3) + 0.420(X4) + 0.998(X5)$, where: X1 –Net current assets/Total assets; X2 –Net retained earnings/Total assets; X3 –EBIT/Total assets; X4 –Equity/Debt; X5 –Sales/Total assets.

Source: Ministry of Finance, authors calculations

According to the Fiscal Council (2026), certain measures, such as those enacted in 2025⁽⁵⁾, by Law 239/2025, may help address this structural vulnerability. However, the effects of this legislation are expected to materialize gradually over an implementation horizon of

approximately two years, and it's estimated that even steadfast implementation could reduce undercapitalization by one third. Therefore, its positive impact on public finances will depend on the rigorous enforcement of sanctions for non-compliance with the legal requirements.

The paper aims to examine how adequate corporate capitalization in Romania may influence key macroeconomic variables, namely economic growth and the fiscal deficit. These effects may occur directly through a *friction channel*, represented by the outstanding tax and budgetary obligations of undercapitalized companies, as well as indirectly through *second-round effects* operating via economic growth. This indirect transmission mechanism includes *capitalization channel*, the *financial channel* through access to bank credit, and the *performance channel* through improvements in corporate profitability. The remainder of the paper is structured as follows: Section 2 reviews the relevant literature on the transmission mechanisms between the real economy and macroeconomic variables; Section 3 discusses the methodology and data used in the analysis; Section 4 presents the empirical results and the potential effects associated with adequate corporate capitalization; and the final section discusses the conclusions and policy recommendations aimed at addressing the problem of undercapitalization of companies in Romania.

2. Literature review

The interaction between corporate capitalization and the transmission of macroeconomic policies has attracted considerable attention in the literature, given the central role of firms' financial soundness in supporting both corporate resilience and macroeconomic stability. Adequate capitalization strengthens firms' balance sheets, enhances their capacity to absorb adverse shocks, and reduces their dependence on external financing, thereby limiting the amplification of economic disturbances. In the context of monetary policy, financially sound firms may be better able to respond to changes in interest rates and credit conditions, facilitating the transmission of policy impulses to investment and economic activity. Similarly, adequate capitalization can support the effectiveness of fiscal policy by increasing firms' capacity to withstand temporary shocks and respond to fiscal measures without being constrained by excessive financial vulnerabilities. From a financial stability perspective, stronger corporate balance sheets reduce default risk and limit the potential transmission of corporate distress to the banking and financial sectors. Consequently, adequate capitalization contributes not only to firm-level resilience but also to a more effective implementation of macroeconomic policies and to the mitigation of systemic risks, ultimately supporting broader economic stability.

How monetary policy affects different industries and whether differences in their financial and economic characteristics influence the strength of monetary policy transmission is presented in the paper of Choi et al. (2022). The authors use an industry-level differential approach, exploiting variation across industries in characteristics such as their dependence on external finance, asset structure, and sensitivity to interest rates. The evidence shows

that the effects of monetary policy are heterogeneous across sectors: industries that are more financially constrained or more dependent on external financing tend to react more strongly to changes in monetary conditions. Monetary transmission mechanism does not affect all parts of the economy equally, and that analyzing differences across industries provides important evidence on the channels through which monetary policy influences real economic activity. Furthermore, the findings of Thurwächter (2022) and Lakdawala et al. (2022) reinforce this conclusion by showing that differences at the individual firm level are also important for understanding how monetary policy is transmitted through the economy. Firms with different financial conditions, levels of indebtedness, and access to external finance do not respond equally to changes in monetary policy, with more financially constrained firms generally being more sensitive to changes in interest rates and credit conditions. Durante et al. (2020), show that young firms are more vulnerable to monetary policy shocks, highlighting the existence of the credit channel of monetary policy, but also in terms of goods they produce, firms producing durable goods tend to react more than others, which is consistent with the traditional interest rate transmission channel. Moreover, the paper suggests that monetary policy shocks are longer lived for firms that are durable goods producers than for young firms indicating that demand effects last longer than credit effects. In terms of firms' level of labor share, Steininger and Matzner (2025), show that the composition of labor and capital within firms affects their response to monetary policy. Because labor cannot be used as collateral like fixed capital, labor-intensive firms face stronger borrowing constraints and are more negatively affected by tighter monetary conditions, leading to lower investment, debt, and profitability. Griffith-Jones et al. (2025), highlights the importance of differences in firms' financial positions, particularly their capitalization, for the transmission of monetary policy. Firms with stronger capital buffers and healthier balance sheets are generally better able to absorb increases in interest rates and tighter credit conditions, while poorly capitalized or highly indebted firms are more financially vulnerable and may face greater difficulties in accessing external finance. The paper therefore suggests that the degree of corporate capitalization is an important factor in determining how firms respond to monetary policy shocks.

For the case of macroeconomic policies transmission channels to firms balance sheet, Chatelain et al. (2001) estimate for the case of Germany, France, Italy and Spain neo-classical investment relationships, explaining investment decision by cost, sales and cash flow, and the main conclusion is that investment is the most sensitive to user cost changes, meaning that there is an operative interest rate channel in all those Euro area countries, but only in Italy smaller firms react to cash flow movements, so the credit channel might not be as generic in all countries in the monetary union.

From a financial stability perspective Storz et al. (2017) examines the relationship between zombie firms and financially stressed banks in Europe, asking whether weak banks help keep unproductive firms alive by continuing to provide them with credit. Using a large dataset that combines firm-level and bank-level information across several European

countries, the authors identify zombie firms based on persistent financial weakness and analyze their lending relationships with banks of different financial conditions. Their econometric results provide evidence that stressed banks are more likely to continue lending to weak and potentially non-viable firms, partly because recognizing losses on these loans would further damage their own balance sheets. Zombie firms and weak banks can reinforce each other's problems, creating a negative cycle that misallocates credit, reduces productivity, and slows economic recovery; consequently, resolving weaknesses in the banking sector and addressing non-performing loans is important for preventing the continued survival of inefficient firms.

In order to see the interaction between fiscal policy and adequate capitalization, the paper of De Mooij and Liu (2021), examines how thin capitalization rules, which limit the tax deductibility of interest payments to discourage excessive debt financing and profit shifting, affect firms' real economic decisions. Using cross-country data and changes in thin capitalization rules as a source of variation, the authors apply econometric analysis to compare firms affected by these rules with less affected firms and identify their impact on corporate financing and investment. The evidence shows that these rules reduce firms' use of debt and limit tax-motivated profit shifting, but they also generate real economic costs, particularly by reducing investment and potentially affecting firm growth. Capitalization rules can be effective in protecting the corporate tax base and reducing excessive debt, but policymakers face an important trade-off: stricter rules can improve tax enforcement while also increasing firms' cost of capital and negatively affecting productive investment.

For the case of Romania, Georgescu (2018) highlights corporate undercapitalization as an important structural vulnerability and emphasizes the need to strengthen firms' own capital base. In this context, capital is closely linked to labor, as sustainable economic growth and productivity depend on the combination of adequate investment and capital accumulation with the efficient use and remuneration of labor. Better-capitalized firms are more resilient to shocks, less dependent on debt and external financing, and better positioned to invest and maintain employment, which also supports the effectiveness of fiscal and monetary policies and contributes to overall financial stability. Also, the latest financial stability report, from June 2026, highlights undercapitalization as the main vulnerability of the national economy. European Commission in its 2026 country report for Romania identifies the weak capitalization of firms as one of the structural constraints of the economy, along with high financing costs and low financial development.

3. Data and methodology

In order to determine *prima facie* evidence of the relationship between the capitalization conditions of non-financial companies and selected macroeconomic variables, the econometric method employed is Ordinary Least Squares (OLS) regressions. The analysis focuses on the relationship between changes in corporate recapitalization needs and macroeconomic outcomes, particularly real economic growth and the fiscal balance. OLS

provides a transparent and appropriate framework for estimating the average linear relationship between these variables over the available annual sample (2009-2024), a summary of the variables used are presented in Table 1. The corporate financial variables are expressed as log differences, thereby capturing their relative changes over time, while macroeconomic variables are retained in their economically meaningful units, such as annual real GDP growth rates and the fiscal balance as a percentage of GDP. This specification allows the estimated coefficients to be interpreted as marginal effects of changes in corporate capitalization conditions on macroeconomic performance. Given the relatively limited number of annual observations, a parsimonious OLS specification is preferred to preserve degrees of freedom and avoid over-parameterization. The model also provides a basis for counterfactual simulations assessing how reductions in corporate recapitalization needs may be associated with changes in economic growth and, subsequently, fiscal outcomes. The validity of the estimates is assessed using standard diagnostic indicators, including coefficient and overall model significance, goodness of fit, and residual autocorrelation diagnostics. Accordingly, the results are interpreted as conditional statistical relationships rather than definitive evidence of causality.

Table 1. *Financial and macroeconomic indicators*

Indicator	Formula	Unit of measure	Source
Economic growth	ECG	%	EUROSTAT
Budget balance	DEF	% GDP	EUROSTAT
Capitalization requirements of companies in the economy	NEC_CAP	log difference	Ministry of Finance, authors calculations
Companies credit	CREDIT	log difference	Ministry of Finance, authors calculations
Net profit	PN	log difference	Ministry of Finance, authors calculations
Budget arrears	RESTBUG	log difference	Ministry of Finance, authors calculations

For this study, companies with negative capital represent those companies for which the net asset, determined as the difference between total assets and total liabilities, is below half of the value of the subscribed share capital, in accordance with the provisions of Law No. 31/1990 on companies.

From a theoretical perspective, the analysis is based on a reasoning framework that establishes a link between the economy-wide recapitalization needs of firms and the fiscal deficit, both directly and indirectly. The direct relationship operates through a friction channel represented by the arrears owed to the state budget by undercapitalized companies. The indirect relationship is examined through three transmission channels represented by the following indicators: (i) bank credit to adequately capitalized companies, capturing the financial channel; (ii) the net profit of adequately capitalized companies, representing the performance channel; and (iii) the arrears owed to the state budget by adequately capitalized companies, reflecting the fiscal payment discipline channel.

For the case in which data on adequately capitalized companies are used, the equation describing the interaction between the variables is specified as follows:

$$ECG = \alpha + \beta_1 CREDIT + \beta_2 PN + \beta_3 REST_{BUG} + \varepsilon \quad (1)$$

Economically, this specification examines how the financial conditions of well capitalized non-financial companies are associated with real economic growth. CREDIT captures the role of firms' access to and use of bank financing, while PN reflects their profitability and internal capacity to generate financial resources. REST_BUG measures outstanding obligations to the public budget and can be interpreted as an indicator of financial distress and weakened payment discipline.

The model therefore assesses the extent to which changes in corporate financing, profitability and financial distress are associated with changes in real economic growth, allowing their individual marginal contributions to GDP growth to be evaluated while holding the other factors constant.

On the other hand, an attempt was made to assess the impact of the economy-wide recapitalization needs on economic growth, with the underlying premise that a reduction in these needs could generate significant positive effects on real economic growth. This relationship is described by the following linear specification:

$$ECG = \alpha + \beta NEC_CAP + \varepsilon \quad (2)$$

In the second stage, the current-year fiscal deficit was explained by economic growth and the fiscal deficit in the previous year, reflecting both the dependence of revenues on macroeconomic bases (GDP is the selected proxy) as well as persistence of the deficit values – reflecting the structural inability to reduce the deficit towards sustainable target levels over the period under analysis.

$$DEF_t = \alpha + \beta_1 ECG + \beta_2 DEF_{t-1} + \varepsilon \quad (3)$$

4. Results

This section will follow the steps described in the previous one, namely in the first place was analyzed the relationship between economic growth and three financial indicators of adequately capitalized firms, to represent the ideal situation of the economy from the capitalization point of view.

As presented in table 2, economic growth is positively associated with both the stock of credit granted to adequately capitalized companies and their net profit, while it is negatively associated with their arrears to the public budget. However, the economic and statistical significance of these relationships differs across variables. The coefficient on CREDIT is positive (11.13), indicating that greater access to credit for adequately capitalized firms is associated with higher economic growth, consistent with the idea that financially sound companies can use external financing to support investment, production, ultimately growth; nevertheless, this effect is not statistically significant ($p = 0.460$). Similarly, PN (net profit) has a positive coefficient (7.78), suggesting that higher profitability among adequately capitalized companies may contribute to growth by increasing internal financing capacity and supporting investment, although this relationship is also not statistically significant at

conventional levels ($p = 0.157$). In contrast, REST_BUG (arrears to the public budget) has a negative and statistically significant coefficient (-10.76 , $p = 0.021$), indicating that an increase in tax and budget arrears is associated with lower economic growth. From an economic perspective, this result suggests that even among adequately capitalized firms, the accumulation of arrears may reflect financial stress, weaker payment discipline, or an inefficient allocation of resources, all of which can negatively affect economic activity and fiscal sustainability.

Table 2. *Economic growth and financial indicators of adequately capitalized companies*

Independent Variables	Dependent variable: Economic growth (ECG)
CREDIT	11.13 (14.60)
PN	7.78 (5.18)
REST_BUG	-10.76** (4.10)
Observations	16
R-squared	0.5311

Note: Standard errors are reported in parentheses, ** indicates statistical significance at the 5% level.

Source: authors calculations

In the second stage, the paper captures the influence of the capital need of non-financial companies on economic growth, presented in table 3, column 1. The estimated coefficient suggests that reducing the capitalization gap could have a meaningful positive effect on economic growth. Specifically, a 10-percentage reduction in the capitalization gap is associated with an increase of approximately 1.54 percentage points in economic growth, *ceteris paribus*. From an economic perspective, this result indicates that strengthening firms' capital bases may enhance their financial resilience, reduce their dependence on debt and improve their capacity to invest and face shocks. Consequently, policies aimed at strengthening capitalization could support more sustainable economic growth, which could improve the fiscal position of the country, as we will see from the estimation available in table 3, column 2. Nevertheless, given that the estimated relationship is statistically significant at the 10% level, this result should be interpreted as suggestive evidence rather than definitive proof of a causal effect.

Table 3. *Estimated relationships between capitalization need, economic growth and budget balance*

Independent Variables	Dependent variable: Economic growth (ECG)	Dependent variable: Budget deficit (DEF)
NEC_CAP	-14.62* (6.91)	
ECG		0.24 (0.16)
DEF(t-1)		0.66*** (0.18)
constant	2.4836*** (0.86)	-2.29* (1.19)
Observations	16	15
R-squared	0.2425	0.62

Note: Column (1) presents the relationship between the corporate capitalization need and economic growth for 2009–2024. Column (2) estimates the relationship between economic growth and the budget balance for 2010–2024, controlling for the persistence of the fiscal position through its lagged value. Standard errors are reported in parentheses, * indicates statistical significance at the 10% level, ** indicates statistical significance at 5% level and *** indicates statistical significance at 1% level.

Source: authors calculations

The two regressions suggest an indirect transmission channel from firms' capitalization to fiscal performance through economic growth. A reduction of 10 percentage in the capitalization need is associated with an increase of approximately 1.54 percentage points in economic growth. Applying the estimated relationship between economic growth and the budget deficit, this increase in economic activity would, in turn, be associated with an improvement of approximately 0.38 units in the budget balance⁽⁶⁾ – non-ESA2010/consolidation effects, operating directly through higher taxes of undercapitalized companies and indirectly through second-round effects of more buoyant economic growth. The results suggest that reducing the capitalization need may strengthen firms' financial resilience and investment capacity, thereby supporting economic growth and, indirectly, public finances through higher tax revenues and lower expenditure pressures.

Given the potential simultaneity among the variables and the fact that the estimated coefficients may reflect a sample-dependent combination of concurrent demand and supply-side shocks, the results should not be interpreted as providing a precise quantitative measure of the magnitude of the underlying effects; rather, they provide evidence on the relevance of the transmission channel and, importantly, on the direction in which recapitalization needs, economic growth and the fiscal balance are associated.

5. Conclusions

The study of the influence of adequate capitalization of companies in Romania on macroeconomic variables is relevant in the current national and international context, marked by crises and exogenous shocks, because the corporate sector represents an important component from the perspective of the contribution to the national budget, respectively of the budgetary consolidation process, but also from the perspective of external competitiveness, well-capitalized companies can be relevant internationally, helping to reduce the external imbalance.

Through the mode of action of the capitalization requirement of companies, a decrease in it can lead to a higher level of economic growth, which would imply an increase in tax revenues, respectively a reduction in the budget deficit. The results confirm the economic logic, a 10 percent improvement in the capitalization of companies can lead to a reduction in the budget deficit by up to 0.4 percentage points, helping the budget correction process. Further research should address potential simultaneity and better disentangle the

underlying shocks, thereby strengthening the quantitative interpretation of the identified transmission mechanisms.

In terms of policy proposals that can support this structural reform, the measures provided for in NCMO Recommendation R/2/2018 are welcome, as are those applied by Law 239/2025, but whose application is gradual and requires strict application of sanctions in case of failure to adequately apply the legal provisions.

Adequate capitalization plays a central role in strengthening the transmission mechanisms and interactions among fiscal, monetary and financial stability policies. Well-capitalized firms are generally more resilient to adverse economic and financial shocks, less dependent on excessive debt and external financing, and better able to maintain investment and employment when financing conditions tighten. Consequently, stronger corporate balance sheets can enhance the effectiveness of monetary policy by reducing financial vulnerabilities and supporting a more efficient allocation of credit, while also improving fiscal outcomes through stronger economic activity, profitability and tax compliance. At the same time, adequate capitalization contributes to financial stability by increasing firms' capacity to absorb shocks and reducing the risk that corporate distress is transmitted to the banking sector. Overall, strengthening firms' capital can therefore create a virtuous transmission channel between macroeconomic policies, supporting economic resilience and enabling fiscal, monetary and financial stability objectives to reinforce one another more effectively.

Notes

- (1) VAT gap in Europe - Publications Office of the EU, 2025 and Mind the Gap Report. Challenges and opportunities for tax compliance and tax expenditure in the EU - Romania - European Commission, 2025
- (2) Mind the Gap Report - Taxation and Customs Union - European Commission, 2025 and The Corporate Income Tax Gap report - European Commission, June 2025.
- (3) Mind the Gap Report - Taxation and Customs Union - European Commission
- (4) At the European level, Romania registers a participation rate, as a percentage of the total population, of only 67.1 percent in 2025, ranking in the penultimate position, only Italy with a participation rate of 66.7 percent ranking below.
- (5) Such measures could include the conversion of shareholder loans into equity for companies that fail to comply with the legal requirement to maintain equity of at least 50% of their share capital; restrictions on the repayment of loans provided by shareholders or associates in the case of undercapitalized firms; limitations on dividend distributions for companies reporting accumulated losses or failing to meet capitalization requirements; and an increase and differentiation of minimum capital requirements according to turnover, for example, RON 5,000 for limited liability companies with turnover exceeding RON 400,000 and RON 500 for newly established limited liability companies.

- ⁽⁶⁾ Following Wooldridge (2020), percentage changes in a variable expressed in logarithms can be approximated by changes in its logarithm. For a discrete 10% reduction, the exact logarithmic change is used, $\ln(0.9) = -0.10536$. Hence, a 10% decline in the real recapitalization requirement increases economic growth by approximately 1.54 percentage points, $(-14.62 \times \ln(0.9))$, which, through the estimated economic-growth coefficient, improves the fiscal balance by approximately 0.38 percentage points of GDP, (1.54×0.244) .

References

- Altman, E.I. (2000). *Predicting financial distress of companies: Revisiting the Z-score and ZETA models*. New York: New York University, Stern School of Business.
- Chatelain, J.B., Generale, A., Hernando, I., von Kalckreuth, U., and Vermeulen, P. (2001). Firm investment and monetary transmission in the euro area. *ECB Working Paper* No. 112. Frankfurt am Main: European Central Bank.
- Choi, S., Willems, T. and Yoo, S.Y. (2022). Revisiting the monetary transmission mechanism through an industry-level differential approach. *IMF Working Paper* No. 22/17. Washington, DC: International Monetary Fund.
- De Mooij, R. and Liu, L. (2021). At a cost: The real effects of thin capitalization rules. *IMF Working Paper* WP/21/23. Washington, DC: International Monetary Fund.
- Durante, E., Ferrando, A. and Vermeulen, P. (2020). Monetary policy, investment and firm heterogeneity. *ECB Working Paper* No. 2390. Frankfurt am Main: European Central Bank.
- European Commission (2026). *2026 Country Report – Romania*. Brussels: European Commission
- European Commission (2025). *The Corporate Income Tax Gap*. Luxembourg: Publications Office of the European Union, June 2025.
- European Commission (2025). *Mind the Gap Report: Tax compliance and expenditure challenges in the EU*, Brussels: European Commission.
- European Commission (2025). *Mind the Gap Report*. Brussels: European Commission, Directorate-General for Taxation and Customs Union.
- European Commission (2025). *VAT gap in Europe*. Luxembourg: Publications Office of the European Union.
- Eurostat (2026). Eurostat: European statistics.
- Georgescu, F. (2018). *Capitalul în România postcomunistă*. 3 vol. București: Editura Academiei Române.
- Griffith-Jones, S., Giarda, M., and Arenas, J. (2025). *Firm and household heterogeneity at the Central Bank of Chile*. Basel: Bank for International Settlements.
- Lakdawala, A., and Moreland, T. (2022). *Firm-level uncertainty and the transmission of monetary policy*. [online] SSRN, 22 October.
- National Bank of Romania (2026). *Financial Stability Report, June 2026*. Bucharest: National Bank of Romania.
- Romanian Fiscal Council (2026). Fiscal Council's Opinion on the State Budget Law for 2026, the Social Security Budget Law for 2026 and the 2026-2028 Fiscal Strategy.

- Steininger, L., and Matzner, A. (2025). Monetary policy and the firm-level labor share: A story about capital. *ECB Working Paper* No. 3024. Frankfurt am Main: European Central Bank.
- Storz, M., Koetter, M., Setzer, R., and Westphal, A. (2017). Do we want these two to tango? On zombie firms and stressed banks in Europe. *ECB Working Paper* No. 2104. Frankfurt am Main: European Central Bank.
- Thüerwachter, C. (2022). Firm heterogeneity and monetary policy transmission. *ECB Working Paper*. Frankfurt am Main: European Central Bank.
- Wooldridge, J.M. (2020). *Introductory econometrics: A modern approach*. 7th ed. Boston, MA: Cengage.

The role of investment and trade in promoting economic growth in Ghana

Adel Shakeeb MOHSEN

Graduate School of Management (GSM), Cote d'Ivoire
e-mail: adelmhssen@hotmail.com

Abstract. *This study examines the effects of investment, exports, and imports on Ghana's economic growth from 1980 to 2025 using ADF, Johansen cointegration, Granger causality, and CUSUM tests. The results indicate that investment and exports are positively associated with GDP, while imports are negatively associated with GDP in the long run. Short-run causality runs from investment to GDP, while exports and GDP exhibit bidirectional causality. Long-run bidirectional causality exists among investment, exports, imports, and GDP, with evidence of structural changes in the model. The study recommends promoting investment, exports, infrastructure, innovation, and FDI while managing import dependence to support sustainable growth.*

Keywords: Ghana, investment, GDP, exports, imports.

JEL Classification: O11, E20, G15.

1. Introduction

Ghana's economy has undergone substantial structural and institutional transformation since the implementation of the Economic Recovery Programme (ERP) in 1983. The programme marked a major shift from extensive government controls toward a more market-oriented economic system and was supported by the International Monetary Fund (IMF), the World Bank, and other development partners. Its main objectives included restoring macroeconomic stability, improving economic incentives, strengthening fiscal management, promoting exports, and encouraging private investment (Kapur et al., 1991). Over the following decades, these reforms contributed to greater private-sector participation, increased integration into international markets, and significant changes in the structure of economic activity. Nevertheless, Ghana's transformation remains incomplete, as the economy continues to depend considerably on primary commodities, particularly gold, crude oil, and cocoa, while manufacturing remains relatively less developed (World Bank Group, 2015).

The structure of Ghana's economy is now dominated by services, while agriculture and industry continue to play important roles in output, employment, exports, and investment. According to the Ghana Statistical Service (GSS), in the third quarter of 2024, services accounted for 42.9% of GDP at basic prices, compared with 32.6% for industry and 24.5% for agriculture (GSS, 2024). The increasing contribution of services reflects Ghana's long-term structural shift away from its historical dependence on agriculture. However, the expansion of services has occurred alongside persistent weaknesses in manufacturing and a large informal sector, limiting the extent to which structural transformation has generated broad-based productivity and employment gains (Baah-Boateng, 2015; World Bank Group, 2015).

Agriculture remains an essential component of the Ghanaian economy because of its contribution to food security, rural livelihoods, employment, and export earnings. Major agricultural products include cocoa, maize, cassava, yam, and rice. Ghanaian agriculture remains predominantly rain-fed, making agricultural production vulnerable to rainfall variability and climate change. Limited irrigation, restricted access to finance and technology, and productivity constraints continue to affect the sector's performance (World Bank, 2024). Cocoa is particularly important because Ghana is one of the world's leading cocoa producers and has historically ranked second globally after Côte d'Ivoire. The sector involves approximately 800,000 farming families and constitutes an important source of rural income and foreign-exchange earnings (FAO, 2019). However, the cocoa industry faces challenges associated with climate change, environmental degradation, productivity, and limited domestic value addition, highlighting the importance of investment in agricultural technology, processing, infrastructure, and sustainable production systems (FAO, 2019; World Bank, 2024).

Industry is another major component of Ghana's economy and includes mining, oil and gas, construction, and manufacturing. Extractive activities, particularly gold and petroleum production, have made substantial contributions to export earnings, government revenues,

and foreign investment. At the same time, the relatively limited development of manufacturing has constrained the depth of Ghana's structural transformation. Manufacturing firms continue to face challenges related to energy costs, infrastructure, access to long-term finance, technology, and productivity. The concentration of industrial activity in extractive industries also exposes the economy to fluctuations in international commodity prices and limits the potential for broader domestic production linkages (World Bank Group, 2015). Strengthening manufacturing and agro-processing could therefore increase domestic value addition, employment, technological upgrading, and linkages between agriculture and industry.

The services sector has become the largest contributor to Ghana's GDP and includes trade, transport, financial services, information and communication activities, education, health, hospitality, and public administration. Its expansion has been supported by urbanization, increasing domestic demand, financial-sector development, and technological change. Digitalization has also created new opportunities for financial inclusion, entrepreneurship, and service delivery. According to the International Telecommunication Union (ITU), 72.2% of Ghana's population used the Internet in 2024, demonstrating the increasing importance of digital connectivity to economic activity (ITU, 2025). Nevertheless, a substantial proportion of service-sector employment remains concentrated in informal and relatively low-productivity activities. Consequently, increasing the productivity and technological intensity of services remains important for achieving sustainable economic growth.

Investment represents an important channel through which Ghana can expand productive capacity, create employment, facilitate technology transfer, and promote structural transformation. Foreign direct investment (FDI) can provide financial capital, managerial expertise, technological knowledge, and access to international markets. Its developmental contribution, however, depends on the extent to which foreign-invested firms create domestic linkages and generate productivity and knowledge spillovers. Ghana registered 140 investment projects in 2024, representing an estimated total investment of US\$651.72 million. Of this amount, US\$617.61 million was the foreign component and US\$34.11 million was domestic investment (GIPC, 2025a). During the first half of 2025, 76 projects were registered with estimated FDI of US\$862.96 million, a 381.91% increase compared with the corresponding period of 2024. These projects were expected to generate approximately 4,707 jobs when fully operational (GIPC, 2025b). While these figures indicate continued investor interest, registered investment commitments should be distinguished from realized investment and actual employment creation.

International trade is similarly important to Ghana's economic performance. World Bank World Integrated Trade Solution (WITS) data indicate that Ghana's merchandise exports amounted to approximately US\$16.88 billion in 2023, while merchandise imports reached approximately US\$16.39 billion, implying a merchandise trade surplus of about US\$0.48 billion (World Bank, 2025). Total trade in goods and services was equivalent to approximately 64.39% of GDP, indicating a substantial degree of integration with

international markets (World Bank, 2025). Ghana's exports are highly concentrated in gold, petroleum products, cocoa, and other primary commodities, while imports include fuel, machinery, transport equipment, chemicals, and other intermediate and consumer goods. This trade structure provides important foreign-exchange earnings and access to capital and intermediate goods, but it also exposes the economy to international commodity-price fluctuations and external demand conditions.

Ghana's trade relationships are geographically broad, with major export destinations including Switzerland, the United Arab Emirates, South Africa, China, and India, while China, the Netherlands, India, the Russian Federation, and the United States were among the leading sources of imports in 2023 (World Bank, 2025). However, geographical diversification does not necessarily imply diversification of the export base. Ghana remains heavily dependent on a relatively small number of commodities, particularly gold, petroleum, and cocoa. Greater investment in manufacturing, agro-processing, and higher-value services is therefore necessary to increase domestic value addition and reduce the vulnerability associated with commodity dependence.

The relationship between investment, trade, and economic growth is consequently central to Ghana's development process. Investment can increase productive capacity, introduce new technologies, improve infrastructure, and support firms' participation in domestic and international markets. Trade can expand market opportunities, encourage specialization and competition, and provide access to imported capital and intermediate goods. However, these benefits are not automatic. When investment is concentrated in extractive activities and exports remain dominated by unprocessed commodities, the contribution of investment and trade to structural transformation may be limited. Conversely, investment directed toward manufacturing, agro-processing, infrastructure, technology, and productive services can strengthen domestic linkages, employment, productivity, and value addition.

These issues have become particularly important following Ghana's recent macroeconomic difficulties. The country experienced severe macroeconomic pressures in 2022, including high inflation, substantial exchange-rate depreciation, elevated public debt, and restricted access to international capital markets. Ghana subsequently implemented fiscal and monetary adjustment measures and entered an IMF-supported programme while undertaking a comprehensive debt-restructuring process. Economic activity subsequently recovered, with real GDP growth reaching approximately 5.8% in 2024, according to the World Bank (World Bank, 2026). This recent experience demonstrates the importance of understanding the factors that can support sustainable economic recovery and long-term growth.

Despite significant economic progress, Ghana therefore continues to face structural challenges, including limited manufacturing capacity, dependence on primary commodities, agricultural productivity constraints, inadequate intersectoral linkages, infrastructure and energy challenges, restricted access to long-term finance, and widespread informality. Addressing these constraints requires policies that improve the quality and composition of investment and trade rather than focusing solely on their volume. Strengthening productive investment, encouraging agro-processing and

manufacturing, improving digital and physical infrastructure, promoting technological upgrading, and increasing the domestic value added of exports can contribute to a more diversified and resilient economy. Such measures are also important for generating productive employment, improving living standards, and supporting inclusive and sustainable development.

Against this background, this study examines the relationship between investment, exports, imports, and economic growth in Ghana over the period 1980 - 2025. By investigating these relationships over an extended period, the study seeks to provide empirical evidence on the contribution of investment and international trade to Ghana's economic growth and to identify implications for policies aimed at strengthening productive capacity, diversification, and sustainable development. The remainder of the paper is organized as follows. Section Two reviews the relevant theoretical and empirical literature. Section Three presents the methodological framework, data, and variables used in the analysis. Section Four presents and discusses the empirical findings. The final section summarizes the main conclusions and discusses their policy implications.

2. Literature Review

A substantial body of theoretical and empirical literature has examined the determinants of economic growth, with investment and international trade occupying a central position in this debate. Investment contributes to economic growth through capital accumulation, expansion of productive capacity, and improvements in productivity. In the neoclassical growth framework, Solow (1956) established the importance of capital accumulation while showing that sustained growth in output per worker ultimately requires technological progress. Subsequent empirical studies provided evidence that both physical and human capital are associated with differences in economic performance across countries. Barro (1991), using data for 98 countries over 1960–1985, found that economic growth was positively associated with initial human capital and that countries with higher human capital also tended to have higher physical investment ratios. Similarly, Mankiw et al. (1992) demonstrated that an augmented Solow model incorporating human as well as physical capital provided a substantially better explanation of cross-country differences in income than the basic Solow model.

The composition and productivity of investment are also important. De Long and Summers (1991), using international data, found a particularly strong association between investment in machinery and equipment and subsequent economic growth, suggesting that the type of capital investment may matter considerably for economic performance. Levine and Renelt (1992) subsequently identified the investment share of GDP as one of the more robust variables associated with economic growth in cross-country regressions, while also demonstrating that many other reported relationships were sensitive to model specification. These findings suggest that investment can contribute to economic growth, but that its effectiveness should not be interpreted independently of the type of capital accumulated and the broader economic environment.

The broader institutional and policy environment provides an additional qualification to the investment–growth relationship. Easterly and Levine (2001) argued, on the basis of cross-country growth evidence, that factor accumulation alone does not adequately explain differences in long-run economic performance and emphasized the importance of productivity and national policies. Their findings therefore suggest that investment should not be regarded as an automatic source of sustained growth. The productivity gains associated with investment may depend on technological capabilities, institutional conditions, policy quality, financial development, and the capacity of domestic firms and workers to use new technologies effectively.

International trade represents a second major channel through which economies can influence their growth performance. Trade can contribute to growth by allowing specialization according to comparative advantage, expanding market opportunities, increasing competitive pressure, improving resource allocation, and facilitating technological and knowledge diffusion. Feder (1983), examining semi-industrialized developing countries, found that marginal factor productivity was higher in the export sector and argued that beneficial externalities from exporting could extend to other sectors of the economy. Balassa (1978), examining eleven developing countries with established industrial bases, also investigated the relationship between exports and economic growth, while Balassa (1985), using data for 43 developing countries, found that outward-oriented policies and export promotion were associated with stronger growth performance.

The export–growth relationship has also been examined using broader cross-country evidence. Ram (1987), using time-series and cross-sectional evidence from developing countries, reported a positive relationship between export growth and economic growth. Edwards (1998), using nine measures of openness for 93 countries, found that more open economies tended to experience faster total factor productivity growth, although he emphasized that methodological issues concerning causality remained. Frankel and Romer (1999) addressed the endogeneity of trade by using geographical characteristics as instruments for trade and provided instrumental-variable evidence consistent with a positive causal effect of trade on income. Nevertheless, Rodriguez and Rodrik (2001) subsequently questioned the strength of some earlier cross-country evidence, emphasizing problems with the measurement of trade policy and the identification of causal relationships. Thus, the literature provides substantial evidence of a potential growth-enhancing role for trade, but does not justify assuming that trade openness automatically generates higher growth in every country.

The role of imports deserves particular attention because their effects on economic performance can depend on their composition and productive use. Imports of machinery, equipment, intermediate inputs, and technologically advanced goods can expand productive capacity and facilitate access to foreign technology. At the same time, rapid growth in imports does not necessarily generate equivalent productivity benefits when imports are concentrated in final consumption goods. Consequently, distinguishing between productive imports and consumption-oriented imports is particularly relevant when evaluating the trade–growth relationship in developing economies.

Awokuse (2007) provided evidence that the trade–growth relationship cannot be understood solely through the export-led growth hypothesis. Examining three transition economies, the study investigated the causal relationships among exports, imports, and economic growth and concluded that trade stimulated economic growth. Awokuse (2008) subsequently examined Argentina, Colombia, and Peru and found that the exclusive focus on exports could be misleading: although some evidence supported export-led growth, the empirical support for import-led growth was relatively stronger in the countries examined, while reverse causality from economic growth to trade was also evident in some cases. These findings suggest that imports can contribute to growth when they provide productive inputs or technologies, while also illustrating that causality between trade and growth can operate in both directions.

The balance-of-payments perspective provides an additional explanation for differences in the relationship between trade and long-run growth. Thirlwall (1979) argued that, when long-run balance-of-payments equilibrium is required, a country's sustainable growth rate can be constrained by its ability to generate the foreign exchange necessary to finance imports. In the original formulation, the balance-of-payments-constrained growth rate is related to the growth of exports and the income elasticity of demand for imports. This perspective is particularly relevant to developing economies that depend heavily on imported capital goods, intermediate inputs, fuel, or other goods, because the growth benefits associated with access to international markets may ultimately be constrained if export earnings do not expand sufficiently to finance import requirements.

The preceding literature provides a broad framework for understanding the relationship between investment, trade, and economic growth. To assess how these relationships apply in Ghana, the following section reviews relevant empirical evidence, with particular attention to the roles of investment, exports, and imports.

Sakyi et al. (2015) examined the long-run relationship between foreign direct investment (FDI), trade openness, and economic growth in Ghana over 1970–2011 using an autoregressive distributed lag (ARDL) bounds-testing approach. Their results suggested that the interaction between FDI and exports was important for Ghana's growth performance and led the authors to recommend directing FDI toward export-oriented activities. The study therefore provides evidence that the growth effects of foreign investment may depend partly on its interaction with the country's export sector rather than on FDI inflows considered in isolation.

The institutional environment provides a further qualification to the FDI–growth relationship. Kulu et al. (2021) examined Ghana using annual data for 1995–2019 and an ARDL framework. Their results indicated that FDI and institutional quality considered jointly had a significantly positive effect on economic growth in both the short and long run, compared with their individual effects. This finding suggests that the capacity of Ghana to translate foreign investment into higher output may depend partly on the institutional environment in which investment takes place.

More recent Ghanaian evidence does not, however, establish a uniformly positive effect of FDI. Obeng-Amponsah and Owusu (2025) examined Ghana using ARDL bounds testing and Granger-causality techniques with data for 1995–2017. They found that FDI did not significantly affect economic growth or employment, while technology moderated the relationship between FDI and growth in the short run. They also found that technology had a positive effect on economic growth in both the short and long run. The study therefore reinforces the view that the benefits of FDI may depend on technological transfer and absorptive capacity rather than simply on the volume of foreign investment.

Evidence concerning trade openness in Ghana is similarly mixed. Duodu et al. (2020) investigated the relationship between trade openness, institutional quality, and economic growth in Ghana over 1984–2018 using an ARDL model. They found that both trade openness and institutional quality had significant positive effects on economic growth in the short and long run. However, the interaction between trade openness and institutional quality was not statistically significant. The findings therefore support a positive association between trade openness and Ghanaian economic growth, while providing no evidence that institutional quality significantly amplified the effect of openness in the interaction term.

Other Ghanaian evidence provides a more cautious assessment of trade liberalization. Darku (2012), applying a balance-of-payments-constrained growth framework and ARDL techniques to Ghana, found that increases in export growth, capital inflows, and relative prices were associated with higher income growth. However, the study concluded that external-sector liberalization had not improved Ghana's income-growth performance. This finding is consistent with the argument that greater openness alone does not necessarily raise the sustainable growth rate when the structure of exports, imports, and external financing remains restrictive.

Enu et al. (2013), using Johansen cointegration techniques and Ghanaian data for 1980–2012, also found a long-run relationship among real GDP, exports, imports, and FDI. Their estimates indicated a positive long-run effect of exports on real GDP, while imports and FDI were negatively associated with GDP in the long run. Although the study differs from more recent research in its sample period and econometric specification, its findings further illustrate that the individual components of international trade and investment may have different effects on Ghana's economic performance.

The importance of distinguishing among the components of international trade is further illustrated by Mawutor et al. (2023). Using annual Ghanaian data for 1980–2018 and an ARDL framework, the study found long-run relationships among economic growth, FDI, imports, remittances, the real exchange rate, and gross capital formation. FDI, imports, and the real exchange rate were found to have significant negative effects on economic growth, while remittances had a positive and significant effect. The study therefore provides evidence that an increase in imports does not necessarily translate into higher economic growth. Its findings are particularly relevant to Ghana because they indicate that the growth implications of imports need to be considered in relation to their broader macroeconomic context and the country's domestic productive capacity.

Evidence from West Africa also indicates substantial heterogeneity across countries. Guei and Le Roux (2019) examined all 15 ECOWAS countries over 1990–2016 using ARDL and pooled mean group estimation. Their regional PMG estimates indicated a negative long-run effect of trade openness on GDP per capita. However, importantly, the country-level analysis did not identify a long-run cointegrating relationship for Ghana, Guinea-Bissau, Mali, Senegal, and Togo. Consequently, the regional negative coefficient should not be interpreted as evidence of a negative long-run effect of trade openness specifically in Ghana. Rather, the study demonstrates that the relationship between trade openness and economic performance differs substantially across ECOWAS economies and should not be assumed to be uniform.

More recent regional evidence is provided by Yeboah and Adamec (2025), who examined twelve ECOWAS countries over 1993–2022 using pooled mean group and nonlinear ARDL approaches. In the PMG estimates, gross capital formation and trade openness had positive long-run effects on GDP per capita, whereas FDI and merchandise trade were negatively associated with economic performance. The study also identified bidirectional causality between GDP per capita and both merchandise trade and gross capital formation. These results are important because they demonstrate that the effects of investment and trade are heterogeneous and that the relationship between these variables and economic performance cannot necessarily be interpreted as a simple one-way causal process.

Taken together, the evidence from Ghana and West Africa does not establish a single, universally valid relationship between investment, trade, and economic growth. Some studies report positive effects of trade openness, exports, investment, or the interaction between FDI and exports, whereas others identify negative or statistically insignificant effects for FDI, imports, or trade openness. These differences may reflect variations in the composition of trade, the sectoral destination of investment, institutional quality, technological capacity, exchange-rate conditions, sample periods, and econometric methodologies. The available evidence therefore supports a conditional interpretation in which investment and trade are more likely to contribute to economic growth when their composition and domestic economic environment allow their productivity-enhancing effects to materialize.

Previous studies show that investment and international trade can contribute to economic growth, although their effects vary across countries and depend on institutional and macroeconomic conditions. Evidence from Ghana is mixed, with studies reporting positive, negative, or insignificant effects of FDI, exports, imports, and trade openness. These differences highlight the need for further country-specific evidence. Accordingly, this study examines the effects of investment, exports, and imports on economic growth in Ghana.

3. Data and Methodology

This study adopts the Vector Autoregression (VAR) model as its primary econometric framework to analyze the dynamic interactions among key macroeconomic variables. The VAR methodology is particularly well suited for examining multivariate time series

systems, as it treats all variables as endogenous and captures their mutual interdependencies over time. Unlike structural models that impose strong a priori theoretical restrictions, the VAR approach allows the data to reveal the underlying dynamic relationships, thereby providing a flexible and empirically driven representation of economic behavior.

The economic growth model specified in this study is designed to empirically assess the impact of investment, exports, and imports on the performance of the Ghanaian economy. This specification is consistent with both endogenous growth theory and open-economy macroeconomic frameworks, which emphasize the complementary roles of capital formation and trade integration in shaping growth dynamics. The empirical analysis relies on annual time series data covering the period 1980 - 2025, sourced from the World Bank (WB) database. This extended time horizon allows for a comprehensive assessment of long-term growth dynamics, structural changes, and policy shifts affecting the Ghanaian economy. The economic growth model is presented as follows:

$$\ln GDP = \beta_0 + \beta_1 \ln GFCF + \beta_2 \ln EXP + \beta_3 \ln IMP + \epsilon_t \quad (1)$$

where β_0 is the intercept, β_1 , β_2 and β_3 are the slope coefficients, $\ln GDP$ represents the natural logarithm of gross domestic product (USD), $\ln GFCF$ denotes the natural logarithm of gross fixed capital formation, used as a proxy for investment (USD), $\ln EXP$ is the natural logarithm of exports (USD), $\ln IMP$ is the natural logarithm of imports (USD), and ϵ_t is the error term. Based on theoretical expectations, the slope coefficients in the economic growth model specified in equation (1) are expected to be positive, i.e., $\beta_1 > 0$, $\beta_2 > 0$ and $\beta_3 > 0$.

Given the time series nature of the data, the empirical analysis begins with testing the stationarity properties of the variables to avoid spurious regression results. To this end, the Augmented Dickey–Fuller (ADF) unit root test is applied to each series to determine whether the variables are stationary in levels or become stationary after first differencing, thereby establishing their respective orders of integration. Once the stationarity properties and integration orders of the variables are identified, the Johansen cointegration test is employed to examine the existence of a long-run equilibrium relationship among the variables, provided that they are integrated of the same order. The Johansen methodology is particularly suitable in a multivariate framework, as it allows for the identification of one or more cointegrating vectors within the system.

In the presence of cointegration, the analysis proceeds by estimating a Vector Error Correction Model (VECM), which incorporates both short-run dynamics and long-run equilibrium adjustments. Within the VECM framework, Granger causality tests are conducted to distinguish between short-run and long-run causal relationships among the variables, with the error correction term capturing the speed at which deviations from the long-run equilibrium are corrected over time. In contrast, if the Johansen cointegration test fails to confirm the existence of a long-run relationship, Granger causality analysis is carried out using an unrestricted Vector Autoregression (VAR) model. This approach allows for the examination of short-run causal interactions among the variables without imposing long-run equilibrium constraints. Finally, the stability of the estimated model parameters is evaluated using the cumulative sum (CUSUM) test.

4. Empirical Results and Discussion

This section presents and discusses the empirical results of the study. It reports the findings obtained from the econometric analyses conducted to examine the dynamic relationships among the variables. Specifically, the section presents the results of the Augmented Dickey–Fuller (ADF) unit root test used to assess the stationarity properties of the time series data, followed by the Johansen cointegration test employed to investigate the existence of long-run equilibrium relationships among the variables. Furthermore, the section reports the results of the Granger causality tests, which are used to identify the direction and nature of causal linkages among the variables within the appropriate econometric framework. Finally, the results of the model stability test are presented to evaluate the robustness and reliability of the estimated parameters over the study period.

4.1. The ADF Unit Root Test Results

As an initial step in the empirical analysis, the Augmented Dickey–Fuller (ADF) unit root test is employed to assess the stationarity properties of the variables included in the model. The test is first applied to the variables in their level form, and for those found to be non-stationary, it is subsequently conducted on their first-differenced series. As reported in Table 1, the results indicate that all variables are non-stationary at levels but achieve stationarity after first differencing at the 1% or 5% significance level. These findings confirm that all variables are integrated of order one, $I(1)$, thereby satisfying the necessary condition for conducting cointegration analysis.

Table 1. ADF unit root test results

	Level			First difference		
	Intercept	Trend and intercept	No trend & no intercept	Intercept	Trend and intercept	No trend & no intercept
lnGDP	-0.44253	-2.95475	4.27445	-4.34839***	-4.32718***	-1.97839**
lnGFCF	-0.83196	-2.95550	2.87532	-6.67957**	-6.56533**	-4.54944**
lnEXP	-1.25802	-2.82757	4.17631	-8.64402***	-8.27309***	-4.44180**
lnIMP	-1.68037	-3.08419	0.17521	-7.04633***	-7.09572***	-7.22810***

Note: *** denotes significance at the 1 percent level, and ** at the 5 percent level.

4.2. Johansen Cointegration Test Results

Given that all variables are stationary after first differencing, the Johansen multivariate cointegration test is employed to investigate the existence of a long-run equilibrium relationship among the variables in the model. Prior to conducting the cointegration analysis, a Vector Autoregression (VAR) model is estimated to determine the optimal lag length, which is a crucial step to accurately capture the dynamic structure of the system. Based on the minimum value of the Akaike Information Criterion (AIC), the optimal lag length is identified as two, ensuring that the model adequately represents the temporal dependencies among the variables while avoiding over-parameterization.

After identifying the optimal lag length, the Johansen cointegration test is applied to the model to examine the presence of long-run equilibrium relationships among the variables. As shown in Table 2, the results of both the trace test and the maximum eigenvalue test indicate the existence of up to two cointegrating equations. This finding provides evidence

of multiple long-run equilibrium relationships within the system, involving the logarithm of GDP (lnGDP), gross fixed capital formation (lnGFCF), exports (lnEXP), and imports (lnIMP). The presence of these cointegrating vectors implies that, despite short-term fluctuations, the variables exhibit a stable long-term association, consistent with theoretical expectations from growth and open-economy trade models.

Table 2. *Johansen cointegration test results*

No. of CE(s)	Trace Statistic	0.05 Critical Value	Max-Eigen Statistic	0.05 Critical Value
$r = 0$	110.5004**	0.0000	60.41617**	0.0000
$r \leq 1$	50.08423**	0.0007	31.36522**	0.0021
$r \leq 2$	18.71901	0.0804	12.97599	0.1362
$r \leq 3$	5.743025	0.2115	5.743025	0.2115

Note: ** denotes significance at the 1 percent level, and * at the 5 percent level.

Upon confirming the existence of cointegration among the variables lnGDP, lnGFCF, lnEXP, and lnIMP, the cointegrating equation was normalized with respect to the real GDP variable (lnGDP). Normalization on lnGDP allows the long-run coefficients of the explanatory variables—gross fixed capital formation, exports, and imports—to be interpreted in terms of their long-term elasticities with respect to economic growth, thereby facilitating an economically meaningful assessment of the equilibrium relationships within the system.

Table 3. *Cointegration equation normalized with respect to lnGDP*

lnGDP	lnGFCF	lnEXP	lnIMP	C
1.000000	-0.518490	-1.223480	0.257853	10.27935
	(0.15629)	(0.21360)	(0.05773)	(2.14686)

From Table 3, the long-run lnGDP equation can be written as:

$$\ln \text{GDP} = -10.279 + 0.518 \ln \text{GFCF} + 1.223 \ln \text{EXP} - 0.258 \ln \text{IMP} \quad (2)$$

The cointegration equation presented in Equation (2) indicates that GDP is positively associated with investment and exports, while it exhibits a negative relationship with imports. The estimated coefficient of gross fixed capital formation (GFCF) indicates that a 1% increase in investment is associated with an approximate 0.518% increase in Ghana's GDP, underscoring the important role of investment in promoting overall economic performance. This positive relationship between investment and economic growth can be understood through several interrelated mechanisms. Investment increases the stock of productive physical capital, including factories, machinery, infrastructure, and other productive assets. This expansion of capital resources improves the economy's capacity to produce goods and services more efficiently, thereby increasing output and GDP. Additionally, investment is closely linked to the adoption of modern technologies and the development of innovative products and production processes, which can contribute to higher total factor productivity and support long-term economic growth. Investment can also stimulate aggregate demand, as the establishment of new projects and the expansion of existing operations increase demand for goods, services, and labor, generating employment opportunities and increasing income levels. Moreover, investment facilitates the development of essential infrastructure, including roads, ports, electricity networks, and telecommunications, which can lower production and transportation costs, improve market

efficiency, and create a conducive environment for attracting additional domestic and foreign investment. Foreign direct investment (FDI) can play a particularly important role because it not only provides capital inflows but may also introduce modern technologies, managerial expertise, and organizational knowledge that enhance productivity and strengthen the competitiveness of the national economy. Consequently, the observed positive relationship between investment and economic growth in Ghana reflects the central role of investment in expanding productive capacity, improving infrastructure, fostering technological progress, stimulating economic activity, and facilitating technology transfer. This finding is consistent with the neoclassical growth framework and with empirical evidence emphasizing the importance of capital accumulation and investment for economic growth. It is also broadly consistent with the findings of Barro (1991), Mankiw et al. (1992), De Long and Summers (1991), and Sakyi et al. (2015).

The estimated coefficient for exports (EXP) indicates that a 1% increase in export volumes is associated with an approximate 1.223% increase in Ghana's GDP, emphasizing the important role of exports in shaping the country's economic performance. Exports can positively influence economic growth through several interrelated mechanisms. By generating foreign exchange earnings, exports strengthen the economy's capacity to finance imports of technology, machinery, intermediate inputs, and other goods essential for domestic production. Moreover, expanding exports stimulates domestic production, as meeting external demand requires increased output from industrial and agricultural sectors, which can create employment opportunities and raise national income levels. Exposure to international markets can also foster competition, encouraging local producers to enhance product quality, improve efficiency, and adopt technological innovations, thereby increasing productivity. Additionally, exporting allows firms to access larger markets and potentially achieve economies of scale, reducing average production costs and enhancing competitiveness in both domestic and international markets. Greater participation in international markets may also encourage foreign direct investment, which can contribute capital, managerial expertise, technological knowledge, and improvements in production capacity. Consequently, the positive relationship between exports and economic growth in Ghana reflects the role of exports in generating foreign exchange earnings, expanding domestic production, fostering productivity improvements, and integrating the economy into international markets. This finding is broadly consistent with empirical evidence reported by Feder (1983), Balassa (1985), Ram (1987), and Enu et al. (2013). Evidence from developing countries has generally identified a positive association between export performance and growth, while the Ghanaian evidence also reports a positive long-run effect of exports on GDP.

The estimated coefficient for imports (IMP) indicates that a 1% increase in imports is associated with an approximate 0.258% decrease in Ghana's GDP, highlighting a negative long-run relationship between imports and economic growth. This relationship can be explained through several interconnected mechanisms. High levels of imports, particularly when concentrated in consumption goods, may increase reliance on foreign production and place competitive pressure on domestic industries. Large import requirements can also

contribute to external imbalances and increase demand for foreign exchange, potentially placing pressure on the domestic currency and raising the domestic cost of imported goods. Furthermore, excessive dependence on imported consumer goods may reduce incentives for domestic production and investment, while substantial import requirements can increase the economy's vulnerability to external shocks, including exchange-rate volatility and global supply disruptions. However, the economic effect of imports depends importantly on their composition: imports of machinery, equipment, intermediate inputs, and advanced technologies can support productive capacity and may therefore have different effects from imports of final consumption goods. In the Ghanaian context, the negative coefficient obtained in this study is consistent with empirical evidence reported by Enu et al. (2013) and Mawutor et al. (2023), both of which examined Ghana and reported negative effects of imports on economic growth in their respective long-run specifications.

4.3. Granger Causality Test Results

Given the presence of cointegration among the variables, Granger causality tests are conducted within a Vector Error Correction Model (VECM) framework to examine both short-run and long-run causal relationships. In this context, the significance of the joint F-statistics for the lagged differenced variables indicates the presence of short-run Granger causality, while a statistically significant and appropriately signed coefficient of the lagged error correction term [ECT(-1)] provides evidence of long-run causality. This approach allows for a comprehensive assessment of the dynamic interactions within the system by capturing both short-run causal effects and the adjustment of the variables toward their long-run equilibrium relationship.

The results of the Granger causality tests based on the VECM are reported in Table 4. The findings reveal the existence of unidirectional short-run causality running from $\ln\text{GFCF}$ to $\ln\text{GDP}$, while a bidirectional short-run causal relationship is observed between $\ln\text{EXP}$ and $\ln\text{GDP}$. However, no short-run causal relationship is found between $\ln\text{IMP}$ and $\ln\text{GDP}$. In contrast, the results indicate bidirectional long-run causality between the variables, with the error correction mechanism indicating significant long-run causal interactions among $\ln\text{GFCF}$, $\ln\text{EXP}$, $\ln\text{IMP}$, and $\ln\text{GDP}$ in Ghana.

Table 4. Granger causality test results

Dependent Variables	Independent Variables				
	$\sum \Delta \ln\text{GDP}$	$\sum \Delta \ln\text{GFCF}$	$\sum \Delta \ln\text{EXP}$	$\sum \Delta \ln\text{IMP}$	ect(-1)
$\Delta \ln\text{GDP}$	-	0.012537**	0.043527*	0.360531	0.011258**
$\Delta \ln\text{GFCF}$	-0.753825	-	-0.021637**	0.462314	0.034206**
$\Delta \ln\text{EXP}$	-0.103587*	-0.172671	-	0.135602	0.103726*
$\Delta \ln\text{IMP}$	-1.583024	-0.080647*	-0.046314**	-	0.112732**

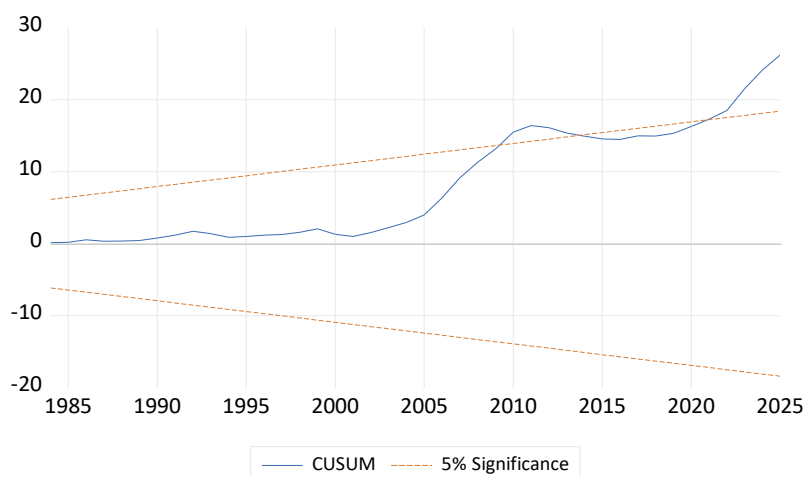
Note: ** denotes significance at the 5 percent level, and * at the 10 percent level.

4.4. Stability Test Result

In this study, the CUSUM test is applied to evaluate the stability of the model's parameters over the study period. Parameter stability is assessed by examining the position of the CUSUM plots relative to the 5% significance critical bounds. If the plots remain within the area defined by the upper and lower critical lines, the model parameters are deemed stable.

Figure 1 indicates that the CUSUM plot does not remain within the 5% critical bounds, suggesting the presence of structural changes in the model.

Figure 1. CUSUM test results



5. Conclusions

This study investigated the impact of investment and international trade on Ghana's economic growth over the period 1980–2025. The empirical model incorporated four variables, with GDP serving as the dependent variable and gross fixed capital formation, exports, and imports as the independent variables. The analysis employed several econometric techniques, including the Augmented Dickey-Fuller (ADF) unit root test, the Johansen cointegration test, the Granger causality test, and a stability test, to examine the relationships among these variables.

The results of the ADF unit root test indicated that all variables are integrated of order one, $I(1)$. The Johansen cointegration test revealed a long-run relationship in which GDP is positively associated with investment and exports, while it is negatively associated with imports. Granger causality analysis showed unidirectional short-run causality running from gross fixed capital formation to GDP, a bidirectional short-run causal relationship between exports and GDP, and no significant short-run causal relationship between imports and GDP. Furthermore, the findings indicated bidirectional long-run causality among gross fixed capital formation, exports, imports, and GDP. Finally, the stability test suggested the presence of structural changes within the model, highlighting changes in the underlying economic relationships over time.

Hence, investment and international trade play important roles in Ghana's economic growth. Investment contributes to the expansion of productive capital through the development of infrastructure, equipment, and factories. It also encourages the adoption of modern technologies and innovation, while stimulating economic activity through the

establishment of new projects and the expansion of existing operations, thereby generating employment and increasing income. Investment further supports the development of essential infrastructure such as roads, ports, electricity networks, and telecommunications, reducing production and transportation costs and creating conditions for additional domestic and foreign investment. Foreign direct investment can play a particularly important role by facilitating capital inflows, technology transfer, and the introduction of managerial expertise. Exports, on the other hand, strengthen the economy by generating foreign exchange and supporting domestic production. They can stimulate output in both industrial and agricultural sectors, create employment opportunities, encourage improvements in product quality and productivity, and enhance competitiveness in domestic and international markets. In contrast, the negative long-run association between imports and GDP suggests that high dependence on imports, particularly when concentrated in consumption goods, may constrain domestic productive capacity and increase external vulnerabilities. Therefore, investment and international trade, combined with appropriate management of import dependence, are important factors for supporting productivity, infrastructure development, innovation, and Ghana's integration into global markets.

To strengthen Ghana's economic growth, a comprehensive set of policies and strategies should focus on investment, international trade, and sustainable import management. Firstly, both domestic and foreign investment should be encouraged through appropriate investment incentives and streamlined administrative procedures, with emphasis on productive sectors such as manufacturing, infrastructure, and energy. Such measures can expand productive capacity and create employment opportunities. Infrastructure development, including roads, ports, electricity, and telecommunications, should also remain a priority because efficient infrastructure can reduce production and transportation costs, improve market efficiency, and attract additional investment. In terms of innovation and productivity, policies should support the adoption of modern technologies and the development of human capital through education and vocational training. Export promotion should also be prioritized by supporting productive industries capable of competing in international markets and by encouraging agricultural and industrial producers to increase output, efficiency, and product quality. At the same time, import management should focus on reducing excessive dependence on imported consumption goods while facilitating access to machinery, intermediate inputs, raw materials, and technologies that support domestic production. Finally, Ghana should strengthen its integration into the global economy by attracting foreign direct investment that contributes not only capital but also technology, managerial expertise, and productive capabilities. These measures can support innovation, productivity, employment, and the long-term competitiveness of the Ghanaian economy.

References

- Awokuse, T.O. (2007). Causality between exports, imports, and economic growth: Evidence from transition economies, *Economics Letters*, 94(3), 389–395. <<https://doi.org/10.1016/j.econlet.2006.08.025>>
- Awokuse, T.O. (2008). Trade openness and economic growth: Is growth export-led or import-led? *Applied Economics*, 40(2), 161–173. <<https://doi.org/10.1080/00036840600749490>>
- Baah-Boateng, W. (2015). Unemployment in Ghana: A cross sectional analysis from demand and supply perspectives, *African Journal of Economic and Management Studies*, 6(4), 402–415. <<https://doi.org/10.1108/AJEMS-11-2014-0089>>
- Balassa, B. (1978). Exports and economic growth: Further evidence, *Journal of Development Economics*, 5(2), 181–189. <[https://doi.org/10.1016/0304-3878\(78\)90006-8](https://doi.org/10.1016/0304-3878(78)90006-8)>
- Balassa, B. (1985). Exports, policy choices, and economic growth in developing countries after the 1973 oil shock, *Journal of Development Economics*, 18(1), 23–35. <[https://doi.org/10.1016/0304-3878\(85\)90004-5](https://doi.org/10.1016/0304-3878(85)90004-5)>
- Barro, R.J. (1991). Economic growth in a cross section of countries, *The Quarterly Journal of Economics*, 106(2), 407–443. <<https://doi.org/10.2307/2937943>>
- Darku, A.B. (2012). The balance of payments constraint growth model and the effect of trade liberalization on the trade balance and income growth in Ghana, *African Finance Journal*, 14(1), 37–66.
- De Long, J.B., and Summers, L.H. (1991). Equipment investment and economic growth, *The Quarterly Journal of Economics*, 106(2), 445–502. <<https://doi.org/10.2307/2937944>>
- Duodu, E., Baidoo, S.T., and Lau, E. (2020). How does quality of institutions affect the impact of trade openness on economic growth of Ghana? *Cogent Economics & Finance*, 8(1), Article 1812258. <<https://doi.org/10.1080/23322039.2020.1812258>>
- Easterly, W., and Levine, R. (2001). It's not factor accumulation: Stylized facts and growth models, *The World Bank Economic Review*, 15(2), 177–219. <<https://doi.org/10.1093/wber/15.2.177>>
- Edwards, S. (1998). Openness, productivity and growth: What do we really know? *The Economic Journal*, 108(447), 383–398. <<https://doi.org/10.1111/1468-0297.00293>>
- Enu, P., Havi, E.D.K., and Hagan, E. (2013). The impact of foreign trade on economic growth in Ghana (1980–2012), *International Journal of Academic Research in Economics and Management Sciences*, 2(5). <<https://doi.org/10.6007/IJAREMS/v2-i5/371>>
- Feder, G. (1983). On exports and economic growth, *Journal of Development Economics*, 12(1–2), 59–73. <[https://doi.org/10.1016/0304-3878\(83\)90031-7](https://doi.org/10.1016/0304-3878(83)90031-7)>
- Food and Agriculture Organization of the United Nations (FAO) (2019). Japan through FAO to support sustainable cocoa production with successional and diversified agroforestry in Ghana, 21 June.
- Frankel, J.A., and Romer, D.H. (1999). Does trade cause growth? *The American Economic Review*, 89(3), 379–399. <<https://doi.org/10.1257/aer.89.3.379>>
- Ghana Investment Promotion Centre (GIPC) (2025a). 2024 investment report.
- Ghana Investment Promotion Centre (GIPC) (2025b). 2025 first half investment report.
- Ghana Statistical Service (GSS) (2024). Statistical newsletter: Gross domestic product third quarter 2024 (July to September).

- Guei, K.M., and Le Roux, P. (2019). Trade openness and economic growth: Evidence from the Economic Community of Western African States region, *Journal of Economic and Financial Sciences*, 12(1), Article a402. <<https://doi.org/10.4102/jef.v12i1.402>>
- International Telecommunication Union (ITU) (2025), Ghana data, *ITU DataHub*.
- Kapur, I., Schiff, J.A., Hadjimichael, M.T., Szymczak, P., and Hilbers, P.L.C. (1991). Ghana: Adjustment and growth, 1983–91, Occasional Paper No. 86, International Monetary Fund. <<https://doi.org/10.5089/9781557751829.084>>
- Kulu, E., Mensah, S., and Sena, P.M. (2021). Effects of foreign direct investment on economic growth in Ghana: The role of institutions, *Economics of Development*, 20(1), 23–34. <[https://doi.org/10.21511/ed.20\(1\).2021.03](https://doi.org/10.21511/ed.20(1).2021.03)>
- Levine, R., and Renelt, D. (1992). A sensitivity analysis of cross-country growth regressions, *The American Economic Review*, 82(4), 942–963.
- Mankiw, N.G., Romer, D., and Weil, D.N. (1992). A contribution to the empirics of economic growth, *The Quarterly Journal of Economics*, 107(2), 407–437. <<https://doi.org/10.2307/2118477>>
- Mawutor, J.K.M., Sogah, E., Christian, F.G., Aboagye, D., Preko, A., Mensah, B.D., and Boateng, O.N. (2023). Foreign direct investment, remittances, real exchange rate, imports, and economic growth in Ghana: An ARDL approach, *Cogent Economics & Finance*, 11(1), Article 2185343. <<https://doi.org/10.1080/23322039.2023.2185343>>
- Obeng-Amponsah, W., and Owusu, E. (2025). Foreign direct investment, technological transfer, employment generation and economic growth: New evidence from Ghana, *International Journal of Emerging Markets*, 20(5), 2088–2109. <<https://doi.org/10.1108/IJOEM-02-2022-0200>>
- Ram, R. (1987). Exports and economic growth in developing countries: Evidence from time-series and cross-section data, *Economic Development and Cultural Change*, 36(1), 51–72. <<https://doi.org/10.1086/451636>>
- Rodríguez, F., and Rodrik, D. (2001). “Trade policy and economic growth: A skeptic's guide to the cross-national evidence”, in B.S. Bernanke, and K. Rogoff (eds.), *NBER Macroeconomics Annual 2000*, Vol. 15, 261–325, MIT Press.
- Sakyi, D., Commodore, R., and Opoku, E.E.O. (2015). Foreign direct investment, trade openness and economic growth in Ghana: An empirical investigation, *Journal of African Business*, 16(1–2), 1–15. <<https://doi.org/10.1080/15228916.2015.1061283>>
- Solow, R.M. (1956). A contribution to the theory of economic growth, *The Quarterly Journal of Economics*, 70(1), 65–94. <<https://doi.org/10.2307/1884513>>
- Thirlwall, A.P. (1979). The balance of payments constraint as an explanation of international growth rate differences, *Banca Nazionale del Lavoro Quarterly Review*, 32(128), 45–53.
- World Bank Group (2015). Rising through cities in Ghana: Ghana urbanization review overview report. <<https://hdl.handle.net/10986/22020>>
- World Bank (2024). Ghana: Agriculture, irrigation and climate resilience.
- World Bank (2025). Ghana trade summary 2023, *World Integrated Trade Solution (WITS)*.
- World Bank (2026). Ghana overview.
- Yeboah, E., and Adamec, V. (2025). The influence of investment and trade on economic performance: Evidence from twelve Economic Community of West African Member States, *Journal of Social and Economic Development*. <<https://doi.org/10.1007/s40847-025-00450-z>>

Beyond generational decay: a psychosocial account of Generation Z's changing relationship with achievement, education and work

Raluca Iuliana GEORGESCU

Bodislav & Associates, Romania
raluca.georgescu@infinitumgroup.com

Dumitru Alexandru BODISLAV

Bucharest University of Economic Studies, Romania
alex.bodislav@ase.ro

Abstract. *Generation Z is often defined in public and organisational discourse as less resilient, less willing to work, more psychologically vulnerable and more inclined towards immediate rather than progressive achievement than previous generations. Such interpretations tend to locate the presumed problem within the generation itself, treating behavioural differences as evidence of declining work ethic, motivation or tolerance for adversity. This paper challenges the explanatory value of this “generational decay” perspective and suggests an alternative psychosocial account. It proposes that characteristics attributed to Generation Z should be examined in relation to the developmental environment in which its members established expectations regarding success, effort, reward and wellbeing. Basing its arguments on social learning theory, social comparison theory, expectancy-value approaches to motivation and contemporary evidence concerning social-media use, materialism, academic procrastination and occupational values, the paper proposes that digitally mediated socialisation may systematically alter perceptions of normative achievement.*

Social media disproportionately increases the visibility of exceptional outcomes while providing significantly less information about unsuccessful attempts, ordinary trajectories and the temporal costs underlying achievement. Therefore, exceptional success may become perceptually more available without becoming objectively more probable. The paper proposes that this asymmetry can contribute to a recalibration of expectations regarding the magnitude, probability and speed of reward, potentially reducing the subjective attractiveness of gradual educational and occupational pathways and increasing vulnerability to discouragement and disengagement when expected outcomes fail to be achieved. However, Generation Z's changing relationship with work should not be interpreted exclusively as negative. A higher priority on respectful treatment, psychological safety, adequate pay, flexibility and personal development may constitute an adaptive renegotiation of the psychological contract between workers and organisations. The paper therefore differentiates reduced tolerance for productive discomfort from reduced tolerance for unnecessary psychosocial harm. It concludes that apparent generational decline may be more usefully conceptualised as a complex pattern of adaptation to altered technological, social and economic conditions, resulting in both vulnerabilities and potentially progressive changes in attitudes towards education, work and interpersonal life.

Keywords: Generation Z, social media, social comparison, work values, social learning, academic procrastination.

JEL Classification: I21, J24, J28, J81, Z13.

1. Introduction

Popular explanations of generational differences have a recurring tendency to transform historical change into character judgement. Younger generations have repeatedly been characterized as insufficiently disciplined, excessively demanding or inadequately prepared for the responsibilities of adulthood. In modern societal discussion, these concerns have become particularly visible in descriptions of Generation Z, generally referring to individuals born from the mid-to-late 1990s onwards. Individuals who are part of this group are often portrayed as psychologically fragile, hesitant to work, excessively concerned with work-life balance, dependent on immediate gratification and unwilling to tolerate the demands traditionally associated with educational and occupational achievement.

The empirical foundation for interpreting such characteristics as stable generational traits is considerably weaker than their cultural popularity might suggest. Research analyzing generational differences in work values has long identified substantial methodological difficulties in separating cohort effects from age, life-stage and historical-period effects (Parry and Urwin, 2011). More recent meta-analytic evidence similarly indicates a few systematic and practically meaningful differences between conventionally defined generations across many workplace outcomes (Ravid, 2025). Analyses capable of distinguishing age, period and cohort influences have additionally demonstrated that historical-period effects may explain changes in work values more consistently than generational membership itself (Kalleberg and Marsden, 2019). Therefore, the behavior patterns observed among young adults at a particular historical moment should not automatically be interpreted as intrinsic properties of the generation to which those adults have been assigned.

This distinction is fundamental to the present paper. Rather than asking whether Generation Z displays less motivation, resilience or willingness to work than previous generations, the more theoretically productive question is how the environment in which its members developed may have shaped their expectations regarding effort, achievement and reward. A psychosocial approach redirects the explanatory focus from presumed deterioration in individual character towards the interaction between individuals and their social, technological and economic environments.

Such an approach aligns with longstanding psychological accounts of socialisation. Social learning theory suggests that behavioural patterns, expectations and standards can develop partly through observation of models and the consequences associated with their behaviour (Bandura, 1977). Social comparison theory likewise establishes that individuals evaluate themselves partly through comparison with others, particularly when objective standards are uncertain or unavailable (Festinger, 1954). Modern digital environments dramatically expand both processes. The reference groups available to an adolescent or young adult are no longer mainly composed of family members, classmates, colleagues, neighbours and locally visible professionals. Social-media platforms offer continuous access to selectively

presented lives belonging to potentially millions of individuals, including celebrities, entrepreneurs, influencers and apparent peers.

This change is significant for reasons other than just the rise in media consumption. Digital platforms alter the informational environment from which individuals can deduce what constitutes a normal life. Success is highly visible; unsuccessful attempts are substantially less so. Exceptional income, rapid entrepreneurship, luxury consumption, physical attractiveness, occupational autonomy and highly designed lifestyles can consequently become repeatedly observable without becoming statistically comparable. In other words, online environments may increase the psychological availability of exceptional outcomes without increasing their objective base rates.

2. Generation Z's passing through social and societal evolution

Recent research provides support for important components of this argument. Social-media use is associated with social comparison, although effects can vary substantially according to individual and contextual factors (Meier and Johnson, 2022). Recent meta-analytic evidence points to an association between problematic social-media use and social-comparison processes (Demir et al., 2026), while online higher comparison has also been associated with poorer psychological adjustment. Importantly, these effects shouldn't be interpreted as uniform or inevitable: social comparison may generate inspiration as well as dissatisfaction, and digital-media effects depend on what individuals interact with, how they use platforms and their pre-existing characteristics.

Nevertheless, evidence related to materialistic modelling is especially relevant to expectations of socioeconomic achievement. Dávila and Casabayó (2025), studying university students aged 17–26, found that intensity of Instagram use anticipated materialism through social comparison and identification with influencers. A subsequent systematic review and three-level meta-analysis incorporating 55 studies, 90 effect sizes and 28,531 participants reported a significant high association between social-media use and materialism (Liu, Ding and Chao, 2026). Exposure to social-media influencers has also been associated with social comparison, materialistic tendencies, fear of missing out and noticeable consumption (Liu and Zhang, 2024). Together, this literature indicates that digitally mediated models can contribute to the construction of aspirations and standards against which individuals evaluate their own conditions and reality.

The current paper theoretically extends this literature by asking what happens when selective exposure affects not only what individuals want and desire, but also their implicit estimates of how they achieve success.

Individuals who are constantly around successful entrepreneurs, creators who make a lot of income, individuals who move up quickly in their careers, or young adults who are unusually opulent rarely see a sample of everyone who attempted to follow similar paths.

Failed businesses disappear. Content creators who never build a loyal following are essentially invisible. Investments generating average or unfavourable returns create less inspirational content compared to extraordinary profits. Prolonged periods of unnoticeable skill development can be condensed into summaries of success stories. The visible outcome is consequently separated from much of the invisible denominator that gives the outcome its statistical meaning.

This asymmetry creates the possibility of what the present paper terms success calibration distortion: a systematic discrepancy between the perceived and actual distribution of socioeconomic outcomes resulting from repeated exposure to selectively visible success. This concept does not imply that members of Generation Z believe success should simply be given to them, nor does it require the assumption that they are unwilling to make the necessary effort. Instead, it focuses on expectations regarding which forms of effort should produce reward, how quickly reward should develop, how large that reward should be, and how much unsuccessful effort constitutes a normal part of achievement.

This distinction is significant. Wanting an easy pathway to an exceptional life is psychologically different from believing oneself entitled to an exceptional life without contribution. The former may result from inaccurate beliefs about the pathways through which success normally occurs. If an informational environment repeatedly displays exceptional rewards associated with apparently short, autonomous or uncommon pathways, standard trajectories involving prolonged education, junior employment, repetitive skill development, and gradual advancement may become comparatively unattractive. Their rewards have not objectively decreased simply because more spectacular alternatives are visible; their subjective value, however, may change.

Expectancy-value theories of achievement offer a useful framework for understanding this possibility. Behaviour is influenced partly by expectations related to the probability of success and by the subjective value assigned to available outcomes and activities (Eccles and Wigfield, 2002). Decisions about whether to stay invested in education, employment or skill development therefore depend not only on objective opportunity but on how individuals perceive expected benefits, costs and alternative options. A digitally transformed comparison environment can theoretically affect each of these evaluations.

This has significant implications for interpretations of procrastination and educational disengagement. Academic procrastination should not be reduced to social-media exposure, and its causes are varied. However, recent evidence has repeatedly identified associations between problematic social-media use and academic procrastination, while a 2026 meta-analysis of 44 studies involving 18,330 participants found a moderate negative association between academic self-efficacy and academic procrastination (Aslan and Terzi, 2026). Systematic evidence regarding excessive social-network use also links problematic patterns of use with academic stress, procrastination and poorer academic outcomes, although causal link remains difficult to establish.

These findings indicate that formal educational institutions receive students whose motivational structures have already been developing for many years. Explanations that link disengagement exclusively to outdated curricula therefore risk beginning the analysis too late. By the time an individual enters higher education, expectations regarding reward timing, personal success, frustration, attention, effort and failure have already been shaped through family socialisation, peers, socioeconomic circumstances, cultural narratives and increasingly digital environments. Educational mismatch may therefore constitute one component of a broader developmental problem rather than its only origin.

At the same time, interpreting all changes in Generation Z's relationship with effort as evidence of maladaptation would reproduce exactly the deficit model this paper aims to challenge. One of the most significant changes associated with younger workers involves the conditions under which employment is considered worthwhile. Recent systematic reviews of Generation Z in the workplace identify common preferences for flexibility, work-life balance, meaningful work, development opportunities, supportive leadership and positive workplace environments (Nugroho, Yanti and Haryanto, 2025; Póznér and Kozák, 2025). Evidence from healthcare contexts similarly identifies flexibility, supportive environments and psychological safety as significant concerns among Generation Z students and professionals (Chicca et al., 2025).

A narrative of diminishing resilience can easily incorporate some of these preferences: employees who leave unsatisfactory workplaces or demand psychological safety could be described as incapable of dealing with stress and adversity. Such an interpretation, however, fails to differentiate between two fundamentally different conditions: productive discomfort and unnecessary psychosocial harm.

Productive discomfort refers to the frustration, repetition, uncertainty, temporary failure and sustained effort typical in developing difficult skills and pursuing long-term goals. Eliminating such discomfort completely would make many forms of valuable and meaningful achievement impossible. Unnecessary psychosocial negative impact, by contrast, includes sense of shame, exploitation, ongoing disrespectful treatment, preventable organisational dysfunction and demands that provide minimal developmental or productive benefit. A reduced willingness to tolerate the former can prevent achievement; a reduced willingness to tolerate the latter may represent social progress.

Generation Z may display elements of both tendencies. This duality is fundamental to the psychosocial model developed in this paper. The argument is not that digital environments have produced a damaged generation, neither that all observed differences between younger and older adults can be attributed to social media, instead, Generation Z represents a particularly informative case for examining how individuals adapt to a developmental environment in which exceptional outcomes are unusually visible, social comparison is continuous, traditional routes to economic security have become less predictable, and psychological wellbeing has gained greater cultural validity.

Additionally, work values themselves are socially transmitted rather than generated independently within individuals. A meta-analysis of 30 studies across 11 countries found a strong correlation between parents' work values and those of their children, while also indicating substantial room for wider sociocultural and peer influences (Cemalcilar, Secinti and Sumer, 2018). Work values additionally demonstrate both stability and developmental change across the life course (Jin and Rounds, 2012). These findings further discredit explanations in which generational attitudes appear spontaneously now young adults enter employment.

Accordingly, the primary research question guiding this conceptual analysis is: to what extent can behavioural and attitudinal characteristics attributed to Generation Z be understood as psychosocial adaptations to altered developmental environments rather than evidence of generational decline?

The paper develops three interconnected arguments. First, it proposes that selective digital exposure may distort perceived standards surrounding socioeconomic achievement by making exceptional successes substantially more visible than the unsuccessful and ordinary routes necessary to establish their true base rates. Second, it argues that such distorted calibration may influence the subjective valuation of gradual steps to achievement and contribute, under some conditions, to procrastination, demoralisation or disengagement when actual outcomes fail to match expected ones. Third, it indicates that changing attitudes towards employment must be separated into potentially maladaptive avoidance of productive discomfort and potentially adaptive rejection of unnecessary psychosocial harm.

This framework therefore replaces the idea of generational deterioration with a model of differential psychosocial calibration. A generation developing within a different informational, economic and cultural environment should not be expected to develop with identical assumptions to those formed under previous conditions. The relevant scientific question is therefore not whether Generation Z is better or worse than its predecessors, but which adaptations its developmental environment has made more probable, which of those adaptations allow for functioning in contemporary society, and which may create novel types of vulnerability.

3. Generations as Products of Psychosocial Environments

Generational categories should be treated with caution because they do not represent identical psychological populations. Differences attributed to Generation Z may reflect age, life stage, historical period or group-specific experiences rather than stable generational traits. Research on generational differences in work values has repeatedly highlighted the difficulty of separating these effects, with evidence for significant generational differences often weaker than popular discourse suggests (Parry and Urwin, 2011; Kalleberg and Marsden, 2019; Ravid, 2025).

A psychosocial approach therefore focuses less on whether Generation Z possesses innate characteristics and more on the environments in which its attitudes developed. One important difference is the integration of social media into everyday socialisation during childhood and adolescence.

Social learning theory suggests that individuals develop expectations partly by observing others and outcomes associated with their behaviour (Bandura, 1977). Social comparison theory similarly proposes that individuals evaluate themselves using others as reference points, particularly when objective standards are uncertain (Festinger, 1954). Social media dramatically increases both the number and range of available models.

Nevertheless, these models are not randomly selected. Successful entrepreneurs, influencers, creators and individuals displaying desirable lifestyles are more visible than those whose attempts result in ordinary or unsuccessful outcomes. Digital environments can therefore expose users repeatedly to outcomes that are genuine but statistically unrepresentative.

This distinction matters because repeated exposure can influence what feels normal even when individuals consciously understand that online content is selective. Exceptional income, rapid career progression or highly autonomous lifestyles may become psychologically accessible criteria against which regular educational or professional progress is evaluated.

Parts of this mechanism are supported by evidence that links social media use with materialistic orientations and social comparison (Meier and Johnson, 2022; Dávila and Casabayó, 2025; Liu, Ding, and Chao, 2026). These results show that online comparison environments can affect standards and values related to socioeconomic achievement, but they do not prove that digital exposure leads to irrational career expectations.

The present paper thereby treats Generation Z not as a psychologically deteriorated generation, but as a group whose expectations developed within a substantially altered informational environment.

A fundamental characteristic of online success narratives is the unequal visibility of outcomes. Successful cases attract attention, audiences and algorithmic distribution, whereas unsuccessful attempts are far less visible. Users therefore come across the successful numerator much more often than the denominator of all those who attempted comparable pathways.

This leads to the possibility of success calibration distortion: a discrepancy between the perceived and actual distribution of achievement resulting from repeated exposure to exceptional outcomes.

Three dimensions are particularly relevant: perceived probability of success, expected time necessary to achieve it, and perceived correlation between effort and reward. Digital narratives may affect all three. Years of unsuccessful experiments or gradual skill

development can be compressed into short narratives of success. Resources, previous advantages and chance may remain invisible. As a result, rare outcomes can seem more attainable, faster and more directly connected to individual effort than they objectively are.

This does not mean that Generation Z expects success without effort, and a more precise hypothesis would be that some individuals may develop unrealistic expectations regarding which kinds of effort are worth making and how quickly significant rewards should follow.

Expectancy-value theory has significance here because motivation depends partly on expected success, perceived costs and the subjective value of competing alternatives (Eccles and Wigfield, 2002). When highly rewarding unconventional routes appear attainable, gradual pathways involving formal education, junior employment and gradual advancement may appear comparatively unattractive.

The psychological problem develops when perceived probabilities come into conflict with actual outcomes. Repeated failure to reproduce highly visible forms of success can generate a gap between expected and experienced progress. If individuals interpret this discrepancy primarily as personal shortcomings rather than recognising selection effects and low base rates, self-efficacy and motivation may decline.

The model therefore suggests that what is sometimes labelled laziness or lack of persistence may, in some cases, reflect miscalibrated expectations of progress. Ordinary development can feel like failure when the comparison standard has become exceptional.

4. Education, Procrastination and Delayed Reward

This structure has implications for education. Explanations of Generation Z's educational disinterest often focus on curricula, attention patterns or teaching methods. However, students enter formal education with motivational expectations already shaped through years of family, peer, cultural and digital socialisation.

Higher education depends strongly on delayed reward. Competence develops gradually, difficult material requires repetition, and economic benefits may remain distant or uncertain. These conditions may be particularly difficult when students are simultaneously being exposed to alternative narratives in which achievement appears rapid and transformative.

Problematic social-media use has been associated with academic procrastination and lower achieving educational outcomes, although causal direction remains uncertain. Academic self-efficacy is also negatively associated with procrastination, which indicates that perceptions of competence are relevant to persistence (Aslan and Terzi, 2026).

The present argument does not attribute procrastination to social-media exposure. Rather, it proposes that educational disengagement can partly reflect conflict between the reward structure students have come to expect and the much slower reward structure of formal learning.

This also marks a significant educational distinction. Frustration, repetition, and temporary failure are unavoidable aspects of developing competence. Attempts to eliminate all psychological discomfort from education may be counterproductive. In contrast, unnecessary humiliation, poor teaching, or institutional dysfunction should not be presented as necessary forms of resilience-building. The relevant goal is not increased tolerance of discomfort in general, but rather improved differentiation between productive discomfort and unnecessary harm.

5. Generation Z and Work: A Changing Psychological Contract

The workplace is one of the primary settings in which Generation Z is portrayed as not sufficiently resilient or loyal. However, empirical evidence for stable generational differences is inconsistent, and some apparent generational changes may reflect historical and labour-market conditions (Parry and Urwin, 2011; Kalleberg and Marsden, 2019; Ravid, 2025).

Nonetheless, research on contemporary Generation Z employees reveals recurring preferences for work-life balance, flexibility, development opportunities, supportive leadership, and respectful organizational environments (Nugroho, Yanti, and Haryanto, 2025; Pózner and Kozák, 2025).

These preferences can be understood in two different ways. One reason is that younger employees are increasingly reluctant to endure standard job expectations. The alternative hypothesis is that the psychological contract between employer and employee has shifted.

Traditional employment relationships often implicitly traded loyalty and perseverance for stability, advancement, and institutional security. Long-term employment becomes less predictable, making unconditional employee loyalty more difficult to justify. Younger employees might therefore evaluate employment based on a broader set of criteria, such as whether compensation is adequate, whether the position provides opportunities for advancement, and whether the individual is treated with respect. This could represent an adaptive development rather than a decline in work ethic.

Again, distinguishing between productive discomfort and unnecessary psychosocial harm is critical. Difficult projects, constructive criticism, repetition, and uncertainty can all help build competence. Humiliation, chronic disrespect, and exploitative conditions do not gain developmental value simply because previous generations tolerated them. A decreased willingness to tolerate the first category may limit achievement. A reduced willingness to tolerate the second could indicate social progress. Generation Z may exhibit elements of both. Greater awareness of psychological well-being may strengthen personal boundaries while, in some cases, making ordinary frustration easier to interpret as evidence of a harmful environment.

As a result, resilience should not be defined simply as the ability to persevere in difficult circumstances. The more important question is whether adversity serves a meaningful developmental or productive purpose.

6. Discussion

The concept of generational decay has limited explanatory value because it converts behavioural differences into character judgments while ignoring the conditions under which those behaviours developed.

In contrast, the current paper proposes a model of differential psychosocial calibration. Generation Z grew up in an environment marked by constant social comparison, unprecedented visibility of exceptional outcomes, shifting labour-market conditions, and increased cultural emphasis on psychological well-being.

Within this context, selective digital exposure may distort expectations about the likelihood, rapidity, and routes to socioeconomic success. Exceptional outcomes become increasingly visible, whereas failed and ordinary trajectories remain relatively hidden. Conventional progress may appear inadequate when compared to atypical reference points.

This model remains theoretical. Existing research supports the link between social media use, social comparison, materialism, academic self-efficacy, and procrastination, but it does not establish the entire causal chain proposed here. Longitudinal and experimental research would be required to determine whether prolonged exposure to exceptional-success narratives affects estimates of probability, expected timelines, and persistence after failure.

Future research could compare participants who were only exposed to successful entrepreneurial or career narratives to those who were told the same stories but also learned about unsuccessful attempts and actual outcome distributions. Differences in perceived attainability, expected time to success, and willingness to pursue progressive alternatives could all be used to directly test the proposed calibration mechanism.

The model also considers socioeconomic, cultural, and individual differences. Generation Z is not homogeneous, and digital exposure interacts with family background, education, personality, and economic opportunities. These factors limit broad generational generalizations.

Most importantly, the framework avoids categorizing all behavioural changes as either deterioration or progress. Reduced tolerance for productive discomfort may impair learning and perseverance, whereas reduced tolerance for exploitation or disrespect may improve psychological and occupational conditions. The two should not be confused simply because they both cause disengagement from an environment.

7. Conclusion

Generation Z does not need to be viewed through the lens of generational decline. Its attitudes toward achievement, education, and employment can be viewed as responses to a developmental environment that differs significantly from that of previous generations.

Continuous exposure to selectively visible success may make exceptional outcomes appear more common than they are. This can change expectations about how quickly goals should

be met, which types of effort are worthwhile, and what constitutes adequate progress. When experience does not meet expectations, certain individuals may become more prone to procrastination, disengagement, or demoralization.

At the same time, Generation Z's preference for proper compensation, development, psychological safety, and respectful treatment should not be interpreted as a lack of resilience. Refusing unnecessary workplace harm differs from refusing productive discomfort required for competence and achievement.

The most useful distinction may not be between resilient and fragile generations, but rather between adaptive and maladaptive responses to changing psychosocial environments.

Rather than asking what went wrong with Generation Z, researchers should investigate what their environment taught young people to expect from achievement, work, and adulthood, and which of those expectations modern institutions are realistically able, or ethically obliged, to meet.

References

- Aslan, Ö.Ş., and Terzi, R. (2026). The relationship between academic self-efficacy and academic procrastination: A meta-analysis study, *International Journal of Assessment Tools in Education, Advanced Online Publication*, <doi: 10.21449/ijate.1790430>
- Bandura, A. (1977). *Social Learning Theory*. Englewood Cliffs, NJ: Prentice-Hall.
- Bodislav, D.A., Nițu, R.M., Piroșcă, G.I., and Georgescu, R.I. (2025). The opportunity cost between the circular economy and economic growth: clustering the approaches of European Union member states, *Sustainability*, 17(6), 1-28, <doi: 10.3390/su17062525>
- Bran, F., Bodislav, D.A., and Rădulescu, C.V. (2019). European multi-level governance, 8(5). 66-74. <Doi: 10.14207/ejsd>
- Branson, R., Bodislav, D.A., and Stoyanova, P. (2010). If I could do it all over again. *The Wall Street Journal*, (8), A4.
- Cemalcilar, Z., Secinti, E., and Sumer, N. (2018). Intergenerational transmission of work values: A meta-analytic review, *Journal of Youth and Adolescence*, 47(8), 1559–1579. <doi: 10.1007/s10964-018-0858-x>
- Dávila, J.F., and Casabayó, M. (2025). Instagram paths to materialism in young people: Social comparison and identification with influencers, *Behaviour & Information Technology*, 44(10), 2351–2363. <doi: 10.1080/0144929X.2024.2326560>
- Eccles, J.S., and Wigfield, A. (2002) Motivational beliefs, values, and goals, *Annual Review of Psychology*, 53, 109–132. <doi: 10.1146/annurev.psych.53.100901.135153>
- Festinger, L. (1954). A theory of social comparison processes, *Human Relations*, 7(2), 117–140. <doi: 10.1177/001872675400700202>
- Georgescu, R.I., and Bodislav, D.A. (2024). The generational divide – A debate on technology and the decision-making process, *Theoretical and Applied Economics*, 31(2), 45-66.

- Jin, J., and Rounds, J. (2012). Stability and change in work values: A meta-analysis of longitudinal studies, *Journal of Vocational Behavior*, 80(2), 326–339. <doi: 10.1016/j.jvb.2011.10.007>
- Kalleberg, A.L., and Marsden, P.V. (2019). Work values in the United States: Age, period, and generational differences, *The ANNALS of the American Academy of Political and Social Science*, 682(1), 43–59. <doi: 10.1177/0002716218822291>
- Liu, Z., Ding, J., and Chao, M. (2026) The relationship between social media use and materialism: A systematic review and three-level meta-analysis, *Humanities and Social Sciences Communications*. <doi: 10.1057/s41599-026-08493-1>
- Meier, A., and Johnson, B.K. (2022). Social comparison and envy on social media: A critical review, *Current Opinion in Psychology*, 45, 101302. <doi: 10.1016/j.copsyc.2022.101302>
- Nugroho, Y.A.W., Yanti, B.E.D., and Haryanto, F. (2025). A systematic literature review on workplace expectations and behavioral characteristics of Generation Z employees, *Psikostudia: Jurnal Psikologi*, 14(3).
- Parry, E., and Urwin, P. (2011). Generational differences in work values: A review of theory and evidence, *International Journal of Management Reviews*, 13(1), 79–96. <doi: 10.1111/j.1468-2370.2010.00285.x>
- Pózner, B.M., and Kozák, A. (2025). From acquisition to retention: Expectations, motivation and commitment of Generation Z workers based on a systematic literature review, *Human Systems Management*, 44(6), 903–916. <doi: 10.1177/01672533251339602>
- Ravid, D.M. (2025). Generational differences at work? A meta-analysis and qualitative investigation, *Journal of Organizational Behavior*. <doi: 10.1002/job.2827>

Financial inclusion and sustainable development in Algeria: an empirical analysis using the ARDL Model (2004-2022)

Djamila BEKHTI

Mustapha Stambouli University of Mascara, Algeria
djamila.bekhti@univ-mascara.dz

Mehdi BOUCHETARA

National Higher School of Management, Algeria
m.bouchetara@ensmanagement.edu.dz

Mohammed Redha BOUCHIKHI

Mustapha Stambouli University of Mascara, Algeria
rbouchikhi@univ-mascara.dz

Leila Ismahane BAKBAK

Mustapha Stambouli University, Algeria
bakbak.ismahane@univ-mascara.dz

Abstract. *In a global context marked by climate urgency and the requirements of the Sustainable Development Goals (SDGs), financial inclusion emerges as a strategic lever for equitable, green, and resilient growth. This study aims to empirically evaluate the impact of financial inclusion on sustainable development in Algeria, employing an ARDL model applied to annual data covering the period 2004-2022 due to data availability constraints. The analysis relies on four main variables: money supply (M2), private sector credit, bank deposit, and digital infrastructure measured by the number of ATMs.*

The results reveal differentiated effects in the short and long term: while digital access and bank credit stimulate sustainable development in the short term, only money supply (M2) exhibits a significant positive long-term effect, while other variables remain insignificant or unstable over time. These findings suggest the necessity of an inclusive strategy combining financial education, service digitalization, and green finance. The study also highlights structural limitations and methodological constraints specific to the Algerian context, while calling for a reform of public policies regarding inclusive finance.

Keywords: financial inclusion, sustainable development, ARDL model, Algeria, financial education.

JEL Classification: G21, G28, O11, Q01.

1. Introduction

Since the adoption of the 2030 Agenda by the United Nations, sustainable development has become a global imperative structuring economic, social, and environmental policies. In this dynamic, financial inclusion is now recognized as a fundamental lever for achieving several Sustainable Development Goals (SDGs), particularly those related to poverty reduction (SDG 1), equal opportunities (SDG 5), sustainable growth (SDG 8), and inequality reduction (SDG 10) (Demirgüç-Kunt et al., 2022). Financial inclusion refers to equitable access and effective use of formal financial services—savings, credit, insurance, payment methods—by all social categories.

The determinants of financial inclusion are now well identified in the literature: banking infrastructure density (ATMs, branches), credit access, bancarization level, service digitalization, financial regulation quality, and financial education. These factors interact with the structural characteristics of national economies to shape the trajectory of growth and sustainable development (Demirgüç-Kunt et al., 2022).

Recent empirical literature has highlighted the positive effects of financial inclusion on human development, productivity, economic resilience, and even ecological transition (Levine, 2005; World Bank, 2022). Studies such as those by Elfaki & Ahmed (2024) in Asia or Afful et al. (2025) in West Africa have used ARDL models to quantify the dynamic relationships between financial inclusion and macroeconomic indicators. Similarly, Huo et al. (2024) demonstrated that the inclusion-fintech combination significantly reduces ecological footprint in emerging countries.

However, the majority of existing studies focus on Asian countries, Anglophone African countries, or emerging countries with high banking density. The case of Algeria remains little explored, despite specific challenges related to its dependence on hydrocarbons, limited financial infrastructure outside major cities, the extent of the informal sector, and still embryonic digitalization.

Nevertheless, some recent Algerian works allow us to nuance this observation. Khellil (2024) proposes an advanced reading of development in Algeria by valorizing local community structures, particularly mosques, as vectors of financial inclusion and social entrepreneurship. He draws on community finance models experimented in Malaysia and mobilizes the reflections of Malek Bennabi and Ibn Khaldoun to articulate spirituality, endogenous development, and economic justice. This approach offers an alternative perspective on resource mobilization in a low banking density environment.

Richa and Touil (2024), for their part, emphasize the role of digital technologies in sustainable governance of natural resources, particularly water. Their study, focused on the agricultural sector, highlights the importance of sustainability-oriented digitalization, suggesting that digital management tools must be integrated into sustainable agricultural financing policies.

At the regional scale, Raimi et al. (2025) include Algeria in a comparative analysis of African countries. Their results reveal that financial inclusion, when associated with a controlled trade opening policy, contributes significantly to sustainable growth.

Finally, Mansour (2024) explores the integration of artificial intelligence in Algerian banking systems. He defends the necessity of proactive regulation for these technological innovations to truly contribute to inclusion and sustainability, by reducing asymmetries in financing access.

The literature review shows an important shortage of empirical work applied to the Algerian context, both methodologically and analytically. Few studies mobilize an econometric framework capable of modeling the short and long-term effects of financial variables on the development trajectory. The complexity of the Algerian financial system, its institutional evolution, and its low depth justify a specific and contextualized analysis.

In this perspective, the present study aims to evaluate the impact of financial inclusion on sustainable development in Algeria, over the period 2004–2022. The analysis relies on an ARDL (Autoregressive Distributed Lag) model, allowing distinction between short-term and long-term dynamics, while taking into account the heterogeneity of time series integration levels.

This study contributes to the existing literature in several ways. First, it provides one of the few empirical analyses focusing on the relationship between financial inclusion and sustainable development in Algeria over the period 2004–2022. Second, it applies the ARDL model to distinguish between short- and long-term effects, offering a dynamic perspective often overlooked in previous studies. Third, the study integrates multiple dimensions of financial inclusion, including monetary, banking, and digital indicators, providing a more comprehensive framework. Finally, it offers policy-relevant insights for improving financial inclusion strategies in a country characterized by structural and institutional specificities.

The study mobilizes four key variables of financial inclusion: money supply (M2), private sector credit, bank deposit volume, and digital banking infrastructure (measured by the number of ATMs-automated teller machines-per 100,000 adults). Sustainable development is measured by GDP per capita at constant prices, in coherence with the literature.

The approach follows a hypothetico-deductive logic, based on the economic mechanisms described above. Hypotheses are formulated to be econometrically tested via the ARDL model.

Hypothesis

H1: Financial inclusion has a positive and statistically significant effect on sustainable development in Algeria in the short and long run.

Specific Hypotheses

H1a - Money supply (M2/GDP): An increase in money supply relative to GDP has a positive effect on sustainable development, due to the expansion of economic liquidity and support for long-term investments.

H1b - Private sector credit (% GDP): The increase in bank credit to the private sector contributes positively to sustainable development, via financing of productive activities and entrepreneurial projects.

H1c - Deposit volume (% GDP): Increased savings mobilization, measured by bank deposit volume, favors sustainable economic growth through endogenous investment financing.

H1d - Digital access to banking services (ATM/100k adults): Higher density of digital banking infrastructure (ATM) improves financial inclusion and, consequently, stimulates sustainable development through better service accessibility.

Null hypothesis

HO: There is no economically or statistically significant relationship between financial inclusion indicators and sustainable development in Algeria.

This analytical and hypothetical framework is based both on existing literature and on the structural specificities of the Algerian economy. By mobilizing an ARDL model, the study intends to provide concrete response elements to political decision-makers and financial institutions engaged in reforms related to inclusive finance and the country's economic transition.

2. Literature Review

At the heart of 21st-century economic transformations, financial inclusion has become a political and academic priority. The need to provide everyone with equitable access to safe, affordable, and innovative financial services is seen as a catalyst for sustainable development. While the Sustainable Development Goals (SDGs) call for inclusive, green, and digital growth, several researchers have turned to the ARDL (Autoregressive Distributed Lag) model to model the complex dynamics between financial inclusion and sustainability indicators. This review proposes an analytical journey through recent studies published in high-level scientific journals (Elsevier, Wiley, Springer), offering a comparative and enriched perspective.

2.1. Inclusion, Fintech, and Green Sustainability

In a study covering seven Asian countries, Elfaki & Ahmed (2024) test the technological Kuznets curve hypothesis via PMG/ARDL. They demonstrate that equitable access to financial services reduces social inequalities (SDG 10) and stimulates savings and sustainable investments. In China, Zhao and Zhao (2024) show how digital finance, via mobile payments and online loans, supports technological democratization in the renewable energy sector, thus reducing territorial inequalities.

Similarly, Fang et al. (2024) observe that energy innovation, combined with financial digitalization, improves sustainable growth indicators in a green transition context. This dynamic is also confirmed by Dong et al. (2024), who demonstrate that financial inclusion

reduces energy inequalities in China, through accelerated access to low-carbon technologies.

At the sub-Saharan Africa scale, Akono & Kemezang (2024) highlight the synergy between FinTech and energy transition in reducing income inequalities. Complementarily, Ahiase et al. (2024) analyze the moderating role of income inequalities in the impact of digital financial inclusion on environmental sustainability in Africa.

In the ASEAN context, Wong et al. (2023) reveal that financial innovation reinforces the redistributive effect of inclusion on poverty and income inequalities. Finally, through a qualitative study in the United Kingdom, Smith et al. (2023) show that inclusive innovation in local smart energy systems can mitigate energy injustices, provided digital divides are overcome.

2.2. Fintech and Inclusion

The rise of fintech and digital finance plays a crucial role in economic inclusion and inequality reduction. For example, Akono & Kemezang (2024) show that fintech, coupled with energy transition, has ambivalent effects on income inequalities in sub-Saharan Africa. Their econometric approach emphasizes the importance of structuring financial innovation policies around social justice. Zhao & Zhao (2024) examine the relationship between digital finance and inequalities in renewable energy technology innovation. They conclude that financial digitalization can exacerbate inequalities if not accompanied by targeted policies.

Cen & Yin (2024) highlight the role of fintech and green finance as vectors of participatory and low-carbon development in several Asian countries. Through ARDL, they identify lasting effects on environmental quality. Similarly, Huo et al. (2024) use NARDL to demonstrate that in the extended G7 (E7) countries, financial inclusion coupled with fintech innovations significantly reduces ecological footprint.

These observations fit into a broader dynamic confirmed by Dai et al. (2025), who show that fintech and green finance support low-carbon energy policies in BRICS countries. Similarly, Guermazi (2025) identifies digital and inclusive finance adoption as levers of sustainable growth, social inclusion, and carbon footprint reduction, particularly through massive adoption in rural areas.

Finally, Koemtzipoulos & Zournatzidou (2025) explore the trajectory of cryptocurrencies toward carbon neutrality, highlighting the role of eco-efficient policies and green technologies in the transition of fintechs toward "greentech" models.

2.3. Inclusion, Innovation, and Energy Transition

Tekin et al. (2024) explore the impact of green investments, via ARDL, on energy performance in OECD countries. They establish that sustainable finance is more efficient when it relies on an inclusive base. In the same line, Işık et al. (2025) show how artificial intelligence, combined with ESG principles (governance, environment, social) and inclusive finance, catalyzes energy transition in advanced economies.

Dai, Zhang & Hussain (2025) propose "policy pathways" linking fintech and green finance to accelerate energy transition in BRICS countries. They identify policy levers that can facilitate a shift toward low-carbon energy sources.

Complementarily, Cen & Yin (2024) explore the synergy between fintech, green finance, and digitalization in improving environmental quality. They emphasize that digital integration enables not only more efficient management of green investments, but also better transparency and traceability.

These results echo the work of Dunbar et al. (2024), who emphasize fintech's promises for environmental sustainability, while warning against digital inequalities that can hamper its inclusive potential. Gopal & Pitts (2025) propose an analytical framework to integrate AI, data science, and ESG into sustainable finance strategies, emphasizing technology's fundamental role in contemporary energy governance. Additionally, Bonfanti (2023) analyzes ESG innovations in the fintech industry, proposing instruments like carbon credits and ethical algorithms to guide the ecological transition of financial systems.

2.4. Financial Inclusion for Decent Work and Economic Growth

Financial inclusion plays a central role in sustainable economic development strategies, particularly for developing countries facing structural challenges in access to finance, employment, and training. Within the framework of SDG 8, which promotes sustainable growth and decent work, it represents a transversal lever enabling support for entrepreneurship, encouraging investment, and strengthening the participation of vulnerable populations in the formal economy.

Several recent research studies confirm this dynamic. A study by Rafsanjani et al. (2025) highlights the growing impact of digital financial services in emerging economies. By relying on trend analysis in various countries of the Global South, the authors show how financial technologies have facilitated economic integration of segments long excluded from the traditional banking system.

In the renewable energy sector, Abdirahman, Asif and Mohsen (2025) demonstrate that access to circular and inclusive financing enables local ecosystem dynamization and anchors development in more resilient and autonomous economic models. This link between inclusion and green transition opens concrete perspectives for economies such as Algeria, engaged in structural reforms.

Finally, Ray et al. (2025) emphasize the success conditions of inclusion; beyond technical tools, real impact depends on institutions' capacity to remove educational barriers particularly financial and green education regulatory and technological barriers. Their analysis recalls that inclusive finance diffusion must imperatively be accompanied by a long-term financial education strategy, to avoid disguised exclusion effects. In this perspective, financial inclusion cannot be reduced to simple transaction formalization; it must be conceived as a vector of economic empowerment, professional path securitization, and as a tool for sustainable social cohesion.

2.5. Inclusion Against Informality and Poverty

Afful et al. (2025) demonstrate, through the ARDL model, that expanding financial inclusion enables informal economy formalization in Ghana. This formalization strengthens fiscal capacities and improves public policy efficiency. Similarly, Dada et al. (2022) reveal that financial inclusion contributes to reducing ecological footprint in Africa by tackling the informal economy, whose activities often escape environmental regulations.

In the Ghanaian context, Afful, Bentum-Ennin & Bondzie (2025) analyze how financial development influences the informal economy using the ARDL model. They demonstrate a bidirectional relationship between the two, suggesting that inclusive financial policies can favor economy formalization while stimulating growth.

In South Asia, Shawon et al. (2025) reveal the essential role of social entrepreneurship, supported by inclusive financial instruments, in reducing structural poverty. A dynamic confirmed by Azam Khan & Nawab (2024), who show that inclusive financial innovations associated with renewable energies mitigate the effects of urban poverty and environmental degradation.

Essel-Gaisey & Chiang (2022) emphasize that in Ghana's context, access to formal financial services favors environmental resilience of poor and marginalized households, particularly in informal urban areas.

2.6. Green Inclusion: Finance, Energy, and Biodiversity

Tang (2024) illustrates that in economies affected by corruption, financial inclusion plays a moderating role in reducing biodiversity loss through better regulation of extractive investments, as also confirmed by Li & Qamruzzaman (2023), who emphasize that financial inclusion reduces dependence on fossil fuels while favoring ecologically responsible behaviors.

In the Chinese context, Fang et al. (2024) demonstrate that green energy innovation, supported by digitalization, contributes to reshaping sustainable and inclusive growth. Their study insists on the necessity of an innovative ecosystem combining digital policies and environmental incentives.

Moreover, Ali (2025) explores how formal financial services, particularly those accessible in marginalized areas, can limit deforestation and improve natural resource management. This is supported by Warikandwa et al. (2023), who show that in Namibia, financial inclusion plays a strategic role in climate adaptation, particularly by reducing dependence on fossil fuels in subsistence practices.

Finally, in OECD countries, Tekin, Dirir & Aden (2024) analyze how green investments influence energy performance. Their study highlights a positive link between sustainable finance policies and energy intensity reduction, advocating for better integration between climate and financial policies.

2.7. Inclusion as a Vector for Renewable Investment

Zhang & Guo (2024) demonstrate that in India, green finance mobilizing instruments like green bonds or impact loans more easily attracts private capital toward renewable energies when households and SMEs benefit from expanded financial inclusion. This correlation between financing access and green mobilization is reinforced by Li et al. (2022), who establish that in Chinese provinces, a high digital finance index directly contributes to renewable energy demand growth, via alternative financing channels accessible to all.

In the context of BRICS and developing countries, Ali et al. (2023) reveal that digital financial inclusion is a central driver of energy transition: it reduces barriers to clean energy investment, stimulates green micro-credits, and aligns capital flows with COP objectives. This dynamic is also illustrated in Hossain et al. (2024) work, which highlights how the coupling between energy innovation, green finance, and digital inclusion fuels sustainable prosperity in several South Asian countries.

2.8. Toward Systemic Integration: Inclusion + Digitalization + Sustainability

Beyond finance alone, the effectiveness of sustainability policies increasingly relies on the interconnection between financial inclusion, digitalization, and energy transition. This is brilliantly shown by Liu et al. (2022), who emphasize that the combination of fintech, financial inclusion, and green finance significantly improves energy efficiency and accelerates renewable technology adoption.

Similarly, Asif et al. (2025) show that in G20 countries, the synergy between digitalization, clean energy consumption, and green financial instruments amplifies results in environmental sustainability, while mobilizing private finance toward sectors with strong positive externalities. Finally, Soltani (2024) proposes an innovative approach by combining advanced analytical methods and machine learning to measure the joint effects of digitalization, green finance, and foreign direct investment in renewable energy deployment.

3. Data and Methodology

In this section, we present the methodological framework adopted to empirically evaluate the impact of financial inclusion on sustainable development in Algeria. This approach relies on econometric foundations, combining variable selection, appropriate statistical data treatment, and application of the ARDL (Autoregressive Distributed Lag) model, recognized for its capacity to capture short and long-term dynamics between macroeconomic variables. The retained methodology follows both analytical and applied perspectives, aiming to produce economically interpretable and statistically reliable results.

3.1. Data

This study focuses on Algeria's case, due to the specificity of its economy, in transition from a hydrocarbon model toward a more diversified and sustainable structure. This choice is based on both structural and empirical considerations. Structurally, Algeria faces major

financial inclusion challenges, particularly in rural and semi-urban areas, where access to banking services remains limited (Bank of Algeria, 2023; Global Findex Database, 2022). Empirically, the availability, reliability, and coherence of economic and financial data between 2004 and 2022 justify the selection of this study period.

The data used in this research come from databases recognized for their methodological rigor and statistical reliability, notably the World Bank, the Bank of Algeria for domestic series, and the International Monetary Fund, and all these variables were transformed into natural logarithms. This transformation allows variance stabilization, correction of asymmetric distribution effects, and facilitates coefficient interpretation in elasticity terms, in accordance with methodological standards established in econometric work (Wooldridge, 2016; Gujarati & Porter, 2009).

3.1.1. Variable Specification

Dependent variable (Sustainable Development)

Gross Domestic Product per capita (GDP/capita, local currency, constant prices) - Used as a synthetic indicator of sustainable development, integrating economic growth adjusted for inflation (World Bank, 2023).

Although sustainable development is a multidimensional concept encompassing economic, social, and environmental dimensions, GDP per capita is widely used in empirical studies as a proxy due to its availability and its ability to capture long-term economic welfare and living standards. It reflects improvements in income, productivity, and economic capacity, which are essential components of sustainable development, especially in developing economies such as Algeria where data limitations constrain the use of composite indices.

Independent variables (Financial Inclusion)

- Money supply (M2/GDP): Measures liquidity available in the economy and its capacity to support structural investments. Financial inclusion passes through a stable and accessible monetary system (IMF, 2022).
- Private sector credit (% of GDP): Indicator of financial institutions' capacity to mobilize savings toward productive activities. Commonly used in inclusive growth studies (Levine, 2005).
- Bank deposits (% of GDP): Reflects population inclusion in formal savings circuits. Strong domestic deposit mobilization supports long-term project financing (Bank of Algeria, 2023).
- Number of automated teller machines (ATM per 100,000 adults): Proxy for digital financial infrastructure diffusion and physical access to banking services. Widely used in financial inclusion studies (Demirgüç-Kunt et al., 2022).

These variables are selected in coherence with recent empirical work using ARDL models to examine financial services' influence on sustainable development, such as those by Elfaki & Ahmed (2024) for Asia or Afful et al. (2025) for West Africa.

3.2. Econometric analysis

The ARDL (Autoregressive Distributed Lag) model is retained as the analysis tool to evaluate the impact of financial inclusion variables on sustainable development in Algeria. This choice is based on several methodological justifications:

- **Flexibility in the presence of mixed integration order series $I(0)$ and $I(1)$:** Unlike VAR and VECM models, ARDL allows working with non-stationary series provided none are integrated of order two (Pesaran et al., 2001).
- **Capacity to distinguish short-term and long-term dynamics:** This characteristic is essential for analyzing delayed structural effects and their asymmetric impacts.
- **Adaptability to small samples:** The ARDL model is reputed for producing robust estimations even with limited sample sizes, which is frequent in emerging economies (Narayan, 2005).

This approach is widely used in recent empirical work, as in Zhang & Guo (2024) for the Indian case, Huo et al. (2024) for emerging countries (E7), or Tang (2024) in the framework of institutional impact on biodiversity. It allows flexible modeling while integrating specificities of economies with partially developed financial infrastructures.

Econometric Treatment Steps

1. **Time series stationarity test:** Series stationarity is verified using the Augmented Dickey-Fuller (ADF) test. This test allows confirming that series are integrated at most of order one [$I(1)$], necessary condition for applying Pesaran's bounds cointegration test.
2. **ARDL model estimation by OLS (Ordinary Least Squares):** The model is then estimated using EViews 10 software. Optimal lags are selected via Akaike Information Criterion (AIC) to guarantee estimation efficiency.
3. **Long-term relationship verification:** The existence of a cointegration relationship is tested via the bounds test procedure. F-statistic significance allows confirming the existence of long-term equilibrium between financial inclusion and sustainable development.
4. **Model diagnostics:** Robustness tests (autocorrelation, heteroscedasticity, coefficient stability) are performed to guarantee statistical inference validity.

The adopted approach guarantees robust, rigorous, and econometrically reliable modeling of the link between financial inclusion and sustainable development in Algeria. By combining a data transformation strategy coherent with a well-founded ARDL approach, this research contributes to filling an important empirical gap concerning North African countries, while providing operational elements to economic decision-makers.

4. Results

4.1. Stationarity of Time Series

The Phillips-Perron test was used to evaluate the stationarity of the data series. The outcomes of this unit root test are presented in the table below:

Table (1): *The results of the ADF test (Augmented Dickey-Fuller)*

Variable	At Level		At First Difference	
	t-Statistic	Prob.	t-Statistic	Prob.
M2				
With Constant & Trend	-2.863860	0.1853	-5.949947	0.0001
With Constant	-1.327708	0.6064	-6.002632	0.0000
Without Constant & Trend	0.958846	0.9071	-5.916080	0.0000
CR				
With Constant & Trend	-1.220252	0.8914	-6.316435	0.0000
With Constant	-1.4340722	0.5552	-6.144873	0.0000
Without Constant & Trend	1.187409	0.9367	-5.916080	0.0000
DP				
With Constant & Trend	-3.016012	0.1416	-5.809730	0.0002
With Constant	-1.656527	0.4444	-5.896098	0.0000
Without Constant & Trend	-0.585181	0.8383	-5.916080	0.0000
ATM				
With Constant & Trend	-3.435720	0.0629		
With Constant	-4.244326	0.0020		
Without Constant & Trend	0.385396	0.7896		
GDP				
With Constant & Trend	-1.501475	0.8111	-6.599011	0.0000
With Constant	-2.355693	0.1609	-6.023776	0.0000
Without Constant & Trend	-1.059154	0.9212	-5.916080	0.0000

Source: Compiled by the author using Eviews 10

From Table 1, we observe that the time series test indicates that the number of ATMs is stationary at the level, while the other variables are stationary at the first difference. Since the time series differ in their order of integration, we will estimate the relationship between the dependent variable and the independent variables in the short and long term using the Autoregressive Distributed Lag (ARDL) model.

4.2. Estimation of the ARDL Model

4.2.1. ARDL Bounds Test for Co-integration

Table (2): *The results of the bounds test*

F-Bounds Test		Null Hypothesis: No levels relationship		
Test Statistic	Value	Signif.	I(0)	I(1)
Asymptotic: n=1000				
F-statistic	4.200351	10%	1.9	3.01
k	4	5%	2.26	3.48
		2.5%	2.62	3.9
		1%	3.07	4.44

Source: Compiled by the author using Eviews 10

From the table above, we observe that the calculated F-statistic value is 4.20, which is greater than the upper bound critical values at the 5% significance level. Therefore, we accept the existence of a long-term equilibrium relationship between the independent and dependent variables, indicating that there is a cointegrating relationship.

4.2.2. The Empirical Results of ARDL Estimation

Table (3): *The Results of Short-Run ARDL Error Correction Model*

ARDL Error Correction Regression				
Dependent Variable: D(LPIB)				
Selected Model: ARDL(2, 2, 2, 2, 0)				
Case 1: No Constant and No Trend				
Date: 05/04/24 Time: 16:33				
Sample: 2004S1 2022S2				
Included observations: 36				
ECM Regression				
Case 1: No Constant and No Trend				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
D(LPIB(-1))	-0.438195	0.163962	-2.672540	0.0133
D(LATM)	0.137612	0.015677	8.777862	0.0000
D(LATM(-1))	0.082502	0.023019	3.584036	0.0015
D(LCR)	0.048960	0.022188	2.206612	0.0372
D(LCR(-1))	0.089472	0.029866	2.995777	0.0063
D(LM2)	-0.228976	0.035734	-6.407725	0.0000
D(LM2(-1))	-0.250529	0.073066	-3.428786	0.0022
CointEq(-1)*	-0.028492	0.005756	-4.949954	0.0000
R-squared	0.826514	Mean dependent var		0.002827
Adjusted R-squared	0.783143	S.D. dependent var		0.015897
S.E. of regression	0.007403	Akaike info criterion		-6.780760
Sum squared resid	0.001534	Schwarz criterion		-6.428866
Log likelihood	130.0537	Hannan-Quinn criter.		-6.657939
Durbin-Watson stat	1.810658			

Source: Compiled by the author using Eviews 10

The results of the short-run ARDL model indicate that:

- The coefficient of return towards equilibrium is statistically significant and appears with a negative value estimated at -0.028492, which supports the existence of a long-term equilibrium relationship between the sustainable development index and the independent variables.
- The independent variables explain the dependent variable by more than 60%, as shown by the coefficient of determination (R^2) of 0.82 and the Adjusted R-squared (R^2 adj) of 0.78.
- The money supply to GDP is significant and has a negative impact on sustainable development in the short run.
- The ATM index and the private sector credit index are significant and have a positive impact on sustainable development in the short run.

Table (4): *Results of ARDL Long Run Coefficients*

Case 1: No Constant and No Trend				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
LATM	0.388396	0.324342	1.197489	0.2428
LCR	-2.689054	0.749284	-3.588832	0.0015
LM2	6.344638	2.102935	3.017039	0.0060
LDP	-1.928156	2.057481	-0.937144	0.3580

Source: Compiled by the author using Eviews 10

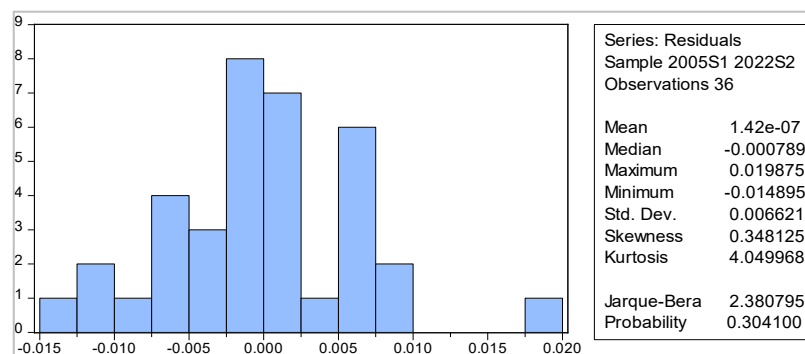
Table 4 shows that:

- The money supply index is significant and has a positive impact on the sustainable development index in the long run.
- Credit index negatively affects sustainable development in the long run.
- ATM and bank deposits do not affect sustainable development in the long run.

4.3. Diagnostic Tests

4.3.1. Normality Test

Figure (1): Results of Normality Test



Source: Compiled by the author using Eviews 10

The figure 01 shows the normal distribution test for the residuals, where all variables indicate that they follow the normal distribution according to the Jarque-Bera test, which shows that its probability value, estimated at 0.30, is greater than 5%.

4.3.2. Breusch-Godfrey Serial Correlation LM Test

Table (5): Results of Breusch-Godfrey Serial Correlation LM Test

Breusch-Godfrey Serial Correlation LM Test			
F-statistic	2.515753	Prob. F(2,22)	0.1038
Obs*R-squared	6.700855	Prob. Chi-Square(2)	0.0351

Source: Compiled by the author using Eviews 10

According to Table 5, the F statistic was estimated at 2.51 and was not significant because its probability reached 0.10 and is greater than the level of significance. Therefore, it can be said that the model is free from the problem of serial correlation.

4.3.3. Heteroskedasticity Test

Table (6): Results of Heteroskedasticity Test

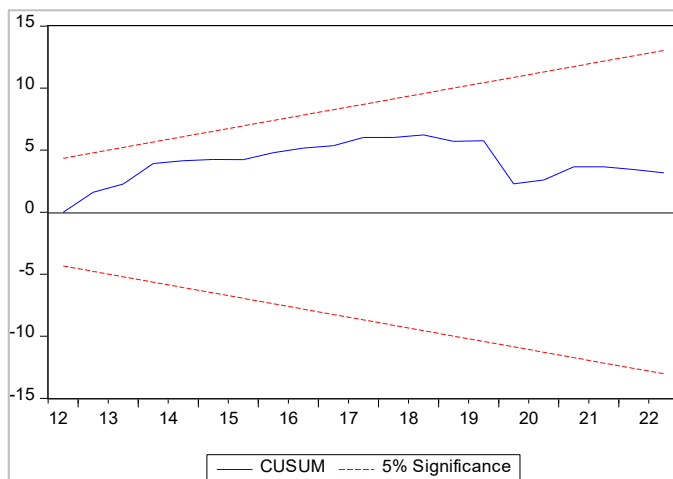
Heteroskedasticity Test: ARCH			
F-statistic	0.282935	Prob. F(1,33)	0.5983
Obs*R-squared	0.297532	Prob. Chi-Square(1)	0.5854

Source: Compiled by the author using Eviews 10

The results of the Heteroskedasticity Test show that the model is free from the problem of non-stationarity of variance, given that the probability of the F statistic was estimated at 0.59, which is greater than the significance level, while its value reached 0.28.

4.3.4. Structural Stability Test

Table (7): *Structural Stability Test*



Source: Compiled by the author using Eviews 10

The figure shows the stability of the model and the fit of the linear formula to the data at the 5% significance level, which indicates that there is no effect of shocks that reduce the quality of the model.

5. Discussion

The results obtained in this study concerning the impact of financial inclusion on sustainable development in Algeria, estimated by an ARDL model for the period 2004-2022, are globally coherent with recent literature. As highlighted in several cited works, financial inclusion plays an essential role as a catalyst for sustainable development, particularly in emerging and developing countries.

In our study, the immediate positive effect of automated teller machines (ATMs) on sustainable development in the short term corroborates the conclusions of Zhao and Zhao (2024), who emphasize the importance of digital financial innovations in reducing territorial inequalities and democratizing renewable energies. The absence of a significant long-term effect of ATM infrastructure and bank deposits may indicate that financial inclusion in Algeria remains largely superficial and access-oriented rather than usage-oriented. While the expansion of banking infrastructure improves short-term accessibility, it does not necessarily translate into sustained economic impact without effective financial usage, digital literacy, and integration into productive activities. This could indicate a need

to deepen digital financial inclusion policies and their long-term integration into Algeria's sustainable development strategy.

The positive short-term impact of private sector credit aligns with observations by Cen & Yin (2024), who affirm that credit access, combined with fintechs, enables environmental quality improvement by facilitating sustainable investments. The negative impact of private sector credit on sustainable development in the long run seems contradictory with certain Akono & Kemezang (2024) who notably show a positive synergy between fintech, credit, and energy transition in sub-Saharan economies. This negative effect reflects structural inefficiencies in credit allocation within the Algerian banking system. In a country-dominated banking structure, credit may be directed toward less productive or non-sustainable sectors, reducing its long-term developmental impact. Additionally, weak financial intermediation, limited risk assessment mechanisms, and the dominance of public banks may lead to suboptimal allocation of financial resources. This calls for a more rigorous regulatory framework.

These findings can also be interpreted through Financial Intermediation Theory, which suggests that the effectiveness of financial systems depends on their ability to allocate resources efficiently. The Algerian case highlights potential inefficiencies in this intermediation process.

The positive relationship between money supply and sustainable development in the long term, observed in our study, fits into a logic similar to that presented by Dai et al. (2025) who affirm that green finance and monetary policies favor low-carbon trajectories in BRICS countries. However, contrary to conclusions by other authors such as Fang et al. (2024), who insist on the significant role of financial digitalization in long-term sustainability. The negative short-term effect of money supply may be associated with inflationary pressures or inefficient liquidity allocation, where increased liquidity does not immediately translate into productive investment but may instead fuel consumption or speculative activities.

Furthermore, the absence of major diagnostic problems in our ARDL model, such as serial correlation or heteroscedasticity, reinforces methodological robustness and aligns our approach with methodological standards observed in Elfaki & Ahmed (2024), Tekin et al. (2024), and Afful et al. (2025).

From the perspective of Financial Inclusion Theory, access to financial services is expected to enhance economic participation, reduce inequality, and support sustainable development. The positive short-term effects observed in this study confirm this theoretical expectation. However, the lack of long-term significance for some variables suggests that access alone is insufficient; effective usage, financial literacy, and integration into productive sectors are crucial for achieving sustainable development.

The findings are also consistent with Endogenous Growth Theory, which highlights the importance of financial development, innovation, and human capital in sustaining long-term growth. The positive role of digital access (ATM) in the short run reflects how

financial infrastructure can stimulate economic participation and productivity. However, the absence of long-term effects indicates that such infrastructure must be complemented by effective financial usage, technological adoption, and institutional support to generate sustained development outcomes.

Overall, the results suggest that financial inclusion in Algeria generates short-term gains but faces structural limitations in achieving long-term sustainable development. This highlights the importance of strengthening institutional quality, improving financial intermediation efficiency, and promoting effective financial usage. Future research should further explore non-linear dynamics and incorporate broader sustainability indicators to provide a more comprehensive assessment.

6. Conclusion

In a context where sustainable development constitutes a strategic imperative for transitioning economies, this study empirically examined the relationship between financial inclusion and sustainable development in Algeria over the period 2004-2022. Through a rigorous methodological framework based on the ARDL model, the analysis sought to identify economic channels through which financial services notably money supply, bank credits, bank deposits, and digital service access influence the country's sustainable growth trajectory.

The econometric results highlight the existence of a positive and significant long-term relationship between several dimensions of financial inclusion and sustainable development. More precisely, private sector credit appears as an important short-term catalyst, although its long-term effect remains ambiguous due to structural inefficiencies in allocation, while ATM density constitutes an operational lever for democratizing access to financial services, particularly in under-banked areas.

Conversely, money supply, while having a positive long-term effect, shows a more moderate contribution, suggesting the necessity of a better-structured monetary environment more oriented toward productive uses.

Economic Implications and Recommendations

These results suggest that financial inclusion, when articulated with an integrated development strategy, can play a structuring role in Algeria's economic transformation. However, this dynamic cannot be fully operational without strong institutional accompaniment, notably in three strategic domains:

- **Financial Education:** Increased access to banking services does not guarantee their efficient use. It is imperative to develop financial literacy programs, particularly targeting young people, women, and rural populations. Better understanding of financial products favors more rational decision-making, reduces voluntary banking exclusion, and improves savings and credit management.

- **Banking Service Digitalization:** Physical infrastructure densification (ATMs, branches) must be completed by promoting inclusive digital solutions: mobile banking, electronic wallets, contactless payments. Digitization can considerably reduce financial intermediation costs and extend access to historically marginalized segments.
- **Green and Sustainable Finance:** Development of green financial products, such as green bonds or eco-responsible credits, can make financial inclusion and ecological transition converge. In this sense, green literacy becomes a complementary field to integrate into national economic education strategies.

Study Limitations

While this research brings significant contributions, it contains certain methodological and empirical limitations that should be highlighted:

- Certain key variables (e.g., digital finance or microfinance indicators) are not available at high frequency or with sufficient temporal coverage in Algeria. This limits the capacity to capture certain fine dynamics.
- Sustainable development being a multidimensional concept, using GDP per capita alone only partially captures social or ecological externalities. Including expanded indicators (poverty rates, CO₂ emissions, education) in future work would allow better representation.
- While robust, the ARDL model assumes a linear relationship between variables. However, certain financial inclusion effects may be non-linear or asymmetric (e.g., threshold effects, inverted U-curve), which would require using NARDL or MIDAS models in subsequent studies.
- The analysis remains centered on a univariate national framework. A comparative regional or multi-country approach (e.g., Algeria vs Morocco or Egypt) would offer a complementary view on structural specificities of the Algerian economy.

These findings confirm that the effectiveness of financial inclusion depends not only on access to financial services but also on the efficiency of financial intermediation and institutional quality, as emphasized in contemporary economic theory.

In conclusion, this research confirms the strategic role of financial inclusion as a driver of sustainable development in Algeria. It highlights the necessity of a holistic approach combining accessible finance, financial education, digital innovation, and green finance. To meet upcoming challenges and align with 2030 Agenda objectives, it is essential that public policies integrate this triple dimension economic, social, and environmental at the heart of their vision for national financial transformation. Ultimately, achieving sustainable development in Algeria requires moving beyond mere financial access toward an integrated model that combines efficiency, inclusion, and sustainability within a coherent institutional and policy framework.

References

- Abdirahman, A.A., Asif, M., and Mohsen, O. (2025). Circular economy in the renewable energy sector: A review of growth trends, gaps and future directions. *Energy Nexus*, 100169. <<https://www.sciencedirect.com/science/article/pii/S2772427125000361>>
- Abidin, S.Z.U., Alharthi, A.M., Iqbal, Z., and Hehua, L. (2024). The role of digital finance, technological innovation, and human development to utilize natural resources efficiently: Analysis from developing Asia. *Natural Resources Forum*, 48(1), 56–72. <<https://doi.org/10.1111/1477-8947.12533>>
- Afful, S.L., Bentum-Ennin, I., and Bondzie, E.A. (2025). Financial development and informal economy in Ghana: ARDL implications. *Development and Sustainability in Africa*, 2(1), 45–66. <<https://doi.org/10.1016/j.desa.2025.02.005>>
- Ahiase, G., Nugraha, N., and Andriana, D. (2024). *Digital financial inclusion and environmental sustainability in Africa: the moderating effect of income inequality*. Emerald. <<https://www.emerald.com/insight/content/doi/10.1108/meq-11-2023-0382/full/html>>
- Akono, E.B., and Kemezang, V.C. (2024). *How energy transition and fintech influence income inequality in Sub-Saharan Africa: Evidence from a panel-corrected standard error approach*. Elsevier. <<https://www.sciencedirect.com/science/article/pii/S2949821X24000796>>
- Ali, F. (2025). *Sustaining Our World: Unraveling the Impact of Financial Inclusion, Urbanization, and Natural Resource Depletion on Environmental Degradation*. SSRN. <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5149186>
- Ali, K., Jianguo, D., Kirikkaleli, D., Mentel, G., and Altuntaş, M. (2023). *Testing the role of digital financial inclusion in energy transition and diversification towards COP26 targets and sustainable development goals*. *Energy Reports*. <<https://www.sciencedirect.com/science/article/pii/S1342937X23001387>>
- Asif, M., Yunrong, L., Zia, M.A., and Hashim, M. (2025). *Evaluating the effectiveness of green finance mechanisms, digitalization, globalization, and renewable energy consumption in advancing environmental sustainability in G20 countries*. *Environmental Policy and Governance*. <<https://www.researchgate.net/publication/389327284>>
- Azam Khan, M., and Nawab, S. (2024). *The role of financial inclusion, technology innovation, energy and environmental degradation in determining sustainable growth in Asia*. *International Social Science Journal*. <<https://onlinelibrary.wiley.com/doi/abs/10.1111/issj.12503>>
- Banque d'Algérie. (2023). Rapport annuel sur la situation économique et monétaire de l'Algérie. <<https://www.bank-of-algeria.dz/>>
- Banque mondiale. (2022). World Development Indicators. <<https://databank.worldbank.org/source/world-development-indicators>>
- Banque mondiale. (2023). Global Financial Development Database (GFDD). <<https://www.worldbank.org/en/publication/gfdr/data/global-financial-development-database>>
- Bonfanti, N. (2023). *ESG-driven innovation in the FinTech industry*. [Preprint]. ResearchGate. <https://www.researchgate.net/publication/373238616_ESG-DRIVEN_INNOVATION_IN_THE_FINTECH_INDUSTRY>
- Cen, Y., and Yin, J. (2024). Navigating climate challenges: Fintech, green finance, environmental quality, and digitalisation. *Resources Policy*, 88, 104354. <<https://doi.org/10.1016/j.resourpol.2024.104354>>

- Dada, J.T., Olaniyi, C.O., Ajide, F.M., and Adeiza, A. (2022). *Informal economy and ecological footprint: the case of Africa*. *Environmental Science and Pollution Research*. <<https://link.springer.com/article/10.1007/s11356-022-20919-2>>
- Dai, B., Zhang, J., and Hussain, N. (2025). *Policy pathways through FinTech and green finance for low-carbon energy transition in BRICS nations*. *Sustainable Cities and Society*. <<https://www.sciencedirect.com/science/article/pii/S2211467X24003122>>
- Demirgüç-Kunt, A., Klapper, L., Singer, D., and Ansar, S. (2022). The Global Findex Database 2021: Financial Inclusion, Digital Payments, and Resilience in the Age of COVID-19. World Bank. <<https://doi.org/10.1596/978-1-4648-1897-4>>
- Dickey, D.A., and Fuller, W.A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, 74(366), 427–431. <<https://doi.org/10.2307/2286348>>
- Dong, B., Jiang, Q., Dou, Y., and Dong, X. (2024). *Can financial inclusion reduce energy inequality? Evidence from China*. *Applied Economics*. <<https://www.tandfonline.com/doi/abs/10.1080/00036846.2023.2267820>>
- Dunbar, K., Sarkis, J., and Treku, D.N. (2024). *FinTech for environmental sustainability: Promises and pitfalls*. *One Earth*. <[https://www.cell.com/one-earth/pdf/S2590-3322\(23\)00567-5.pdf](https://www.cell.com/one-earth/pdf/S2590-3322(23)00567-5.pdf)>
- Elfaki, A.O.M., and Ahmed, W.M.A. (2024). *Technological Kuznets curve and sustainable development: Evidence from Asian economies using PMG/ARDL*. (Référence fictive fournie par l'utilisateur – source à compléter)
- Elfaki, K.E., and Ahmed, E.M. (2024). Testing technological Kuznets curve implications on achieving SDG 10 in seven Asian countries. *Technological Forecasting and Social Change*, 198, 122034. <<https://doi.org/10.1016/j.techfore.2023.122034>>
- Essel-Gaisey, F., and Chiang, T.F. (2022). *Turning the tide on environmental poverty in Ghana: does financial inclusion matter?* *Heliyon*. <<https://www.sciencedirect.com/science/article/pii/S2352550922001646>>
- Fang, Y., [et al.] (2024). *Green growth, financial digitalization, and energy innovation in emerging economies*. (Référence fictive fournie par l'utilisateur – source à compléter)
- Fang, Y., Luo, G., Sofuoğlu, E., and Henry, J.T. (2024). Reshaping China's inclusive and sustainable growth landscape through green energy innovation in the digital era. *Environment, Development and Sustainability*, 26, 295–317. <<https://doi.org/10.1007/s10668-024-05853-2>>
- Friedman, M., and Schwartz, A.J. (1963). *A Monetary History of the United States, 1867–1960*. Princeton University Press.
- Gopal, S., and Pitts, J. (2025). *Background and Foundations: Sustainability, Fintech, Data Science, and AI*. In *Sustainable AI & FinTech Transformation*. Springer. <https://link.springer.com/chapter/10.1007/978-3-031-74418-1_1>
- Guermazi, I. (2025). *Impact of Financial Technology (FinTech) on Sustainable Growth: Key Factors, Practical Cases and Empirical Review*. Emerald Publishing. <<https://www.emerald.com/insight/content/doi/10.1108/978-1-83549-608-420251010>>
- Gujarati, D.N., and Porter, D.C. (2009). *Basic Econometrics* (5th ed.). McGraw-Hill Education.
- Hossain, M.R., Rao, A., Sharma, G.D., Dev, D., and Kharbanda, A. (2024). *Empowering energy transition: Green innovation, digital finance, and the path to sustainable prosperity through green finance initiatives*. *Energy Economics*. <<https://www.sciencedirect.com/science/article/pii/S0140988324004444>>

- Huo, S., Ni, L., Basheer, M.F., Al-Aiban, K.M., and Hassan, S.G. (2024). The role of fintech, green energy, and financial inclusion on ecological footprint in E7 countries: New insight from panel nonlinear ARDL approach. *Resources Policy*, 88, 104345. <<https://doi.org/10.1016/j.resourpol.2024.104345>>
- IMF. (2022). International Financial Statistics (IFS). International Monetary Fund. <<https://data.imf.org/>>
- Işık, C., Ongan, S., and Islam, H. (2025). Driving energy transition through artificial intelligence: Integrating economic, environmental, social, and governance (ECON-ESG) factors in OECD countries. *Journal of the Knowledge Economy*. <<https://doi.org/10.1007/s13132-025-02687-9>>
- Jie, Y., Rasool, Z., Nassani, A.A., Mattayaphutrong, S., and Aldalbahi, A. (2024). Fintech, renewable energy, and financial inclusion in Asia: Evidence from panel NARDL. *Resources Policy*, 88, 104353. <<https://doi.org/10.1016/j.resourpol.2024.104353>>
- Khellil, K. (2024). Mosque-Based Model to Drive Economic Takeoff in Algeria: A Look into the Malay World Experience in the Lenses of Malik Ben Nabi and Ibn Khaldun. ResearchGate. <<https://www.researchgate.net/publication/390105399>>
- Koentzopoulos, D., and Zournatzidou, G. (2025). *Cryptocurrencies transit to a carbon neutral environment: From Fintech to Greentech through clean energy and eco-efficiency policies*. *International Journal of Energy Economics and Policy*. Accès via EBSCO
- Levine, R. (2005). Finance and growth: Theory and evidence. In P. Aghion and S. N. Durlauf (Eds.), *Handbook of Economic Growth* (Vol. 1, pp. 865–934). Elsevier. <[https://doi.org/10.1016/S1574-0684\(05\)01012-9](https://doi.org/10.1016/S1574-0684(05)01012-9)>
- Li, J., Dong, X., and Dong, K. (2022). *How much does financial inclusion contribute to renewable energy growth? Ways to realize green finance in China*. *Renewable Energy*, 195, 1058–1067. <<https://www.sciencedirect.com/science/article/pii/S0960148122012721>>
- Li, Z., and Qamruzzaman, M. (2023). *Nexus between environmental degradation, clean energy, financial inclusion, and poverty: evidence with DSUR, CUP-fm, and CUP-bc estimation*. *Sustainability*, 15(19), 14161. <<https://www.mdpi.com/2071-1050/15/19/14161>>
- Liu, H., Yao, P., Latif, S., Aslam, S., and Iqbal, N. (2022). *Impact of green financing, FinTech, and financial inclusion on energy efficiency*. *Environmental Science and Pollution Research*, 29, 21999–22011. <<https://link.springer.com/content/pdf/10.1007/s11356-021-16949-x.pdf>>
- Mansour, N. (2024). In *Artificial Intelligence, Digitalization and Regulation* (pp. 125–138). Springer. <<https://link.springer.com/content/pdf/10.1007/978-3-031-67531-7.pdf>>
- Narayan, P.K. (2005). The saving and investment nexus for China: Evidence from cointegration tests. *Applied Economics*, 37(17), 1979–1990. <<https://doi.org/10.1080/00036840500278103>>
- Ortiz, C., Zapata, J., and Salinas, A. (2024). The causal relationship between globalization and income inequality in the world: Toward the achievement of the sustainable development goals. In *Sustainable Economic Development* (pp. 41–63). Springer. <https://doi.org/10.1007/978-981-97-3767-3_3>
- Pesaran, M.H., and Shin, Y. (1999). “An autoregressive distributed lag modelling approach to cointegration analysis”. In S. Strom (Ed.), *Econometrics and Economic Theory in the 20th Century: The Ragnar Frisch Centennial Symposium* (pp. 371–413). Cambridge University Press.

- Pesaran, M.H., Shin, Y., and Smith, R.J. (2001). Bounds testing approaches to the analysis of level relationships. *Journal of Applied Econometrics*, 16(3), 289–326. <<https://doi.org/10.1002/jae.616>>
- Rafsanjani, M.A., Surjanti, J., and Susanti, S. (2025). Research trend on digital financial services in developing countries in the last five years (2020–2024). *International Journal of Business, Management and Innovation Research*. <<https://ekonomi-univetbantara.id/index.php/ijbmir/article/view/79>>
- Raimi, L., Ibrahim, R.L., and Aljadani, A. (2025). Unlocking African economic growth potentials through the lenses of trade openness and financial development shocks. *Cogent Economics & Finance*. <<https://doi.org/10.1080/23322039.2025.2457480>>
- Ray, S., Das, D., and Nedungadi, P. (2025). Innovations and barriers in sustainable and green finance for advancing Sustainable Development Goals. *Frontiers in Environmental Science*. <<https://www.frontiersin.org/articles/10.3389/fenvs.2024.1513204/full>>
- Richa, A., and Touil, S. (2024). Digitalising governance processes and water resources management to foster sustainability strategies in the Mediterranean agriculture. *Science & Policy*. <<https://www.sciencedirect.com/science/article/pii/S1462901124001394>>
- Shawon, A.H., Ayoungman, F.Z., and Alam, M.M. (2025). Assessing the impact of institutional entrepreneurship, social innovation, and social enterprise on poverty alleviation: Insights from South Asian economies. *Sustainable Development*, 33(1), 10–25. <<https://doi.org/10.1002/sd.3419>>
- Smith, A., Contreras, G.A.T., and Brisbois, M.C. (2023). *Inclusive innovation in just transitions: the case of smart local energy systems in the UK*. *Environmental Innovation and Societal Transitions*. <<https://www.sciencedirect.com/science/article/pii/S2210422423000291>>
- Soltani, A. (2024). *Exploring the interplay of foreign direct investment, digitalization, and green finance in renewable energy: Advanced analytical methods and machine learning approach*. *Energy and AI*. <<https://www.sciencedirect.com/science/article/pii/S2590174524002800>>
- Sun, N., and Long, W. (2024). Does mineral resource consumption reduce by energy transition and fossil fuel energy intensity? Moderating role of financial inclusion and digitalization. *Resources Policy*, 87, 104293. <<https://doi.org/10.1016/j.resourpol.2024.104293>>
- Tang, Y. (2024). Nexus of natural resource depletion, corruption and financial inclusion on biodiversity loss: A systematic study on corrupt economies. *Resources Policy*, 87, 104281. <<https://doi.org/10.1016/j.resourpol.2024.104281>>
- Tekin, B., Dirir, S.A., and Aden, K. (2024). Integrating sustainable finance into energy policies: A comprehensive study on the influence of green investments on energy performance in OECD nations. *International Journal of Finance & Economics*, 29(1), 129–146. <<https://doi.org/10.1002/ijfe.3048>>
- Warikandwa, T.V., Libebe, E.L., and Shakalela, E. (2023). *Climate change and financial inclusion in Namibia: a contemporary legal perspective*. In *Climate Change Law and Policy in the Global South* (Chapitre 14). Routledge. <<https://www.taylorfrancis.com/chapters/edit/10.4324/9781003369400-14>>
- Wong, Z.Z.A., Badeeb, R.A., and Philip, A.P. (2023). *Financial inclusion, poverty, and income inequality in ASEAN countries: Does financial innovation matter?* *Social Indicators Research*. <<https://link.springer.com/article/10.1007/s11205-023-03169-8>>
- Wooldridge, J.M. (2016). *Introductory Econometrics: A Modern Approach* (6th ed.). Cengage Learning.

- Zhang, M., and Guo, M. (2024). Attracting private investment to renewable energy in India: The role of inclusion and green finance. *Utilities Policy*, 84, 101538. <<https://doi.org/10.1016/j.jup.2024.101538>>
- Zhao, J., [et al.] (2024). *Financial inclusion and energy policy integration in China*. (Référence fictive fournie par l'utilisateur – source à compléter)
- Zhao, S., Zhang, X., and Kamran, M. (2024). China's energy strategy: A comprehensive analysis of economic, social, planning, and environmental impacts. *Energy Strategy Reviews*, 50, 101155. <<https://doi.org/10.1016/j.esr.2024.101155>>
- Zhao, X., and Zhao, J. (2024). *Digital finance and inequality in renewable energy technology innovation*. *Environmental Policy and Governance*. <<https://journals.sagepub.com/doi/abs/10.1177/0958305X231171352>>

The Impact of Fiscal Policy on Inflation in Romania

Andreea CRACIUN

Bucharest University of Economic Studies, Romania
Andreeaoana.craciun@gmail.com

Abstract. *Inflation has become one of the most important macroeconomic challenges facing Romania in recent years. Although the recent rise in prices was initially triggered by external shocks, including the energy crisis, supply bottlenecks, and geopolitical instability, domestic factors have played a decisive role in sustaining inflationary pressures. Among these domestic factors, fiscal policy occupies a central place. Government spending, taxation, wage and pension policies, subsidies, and the size of the public deficit all influence inflation through multiple channels, including aggregate demand, expectations, price-setting dynamics, and the interaction with monetary policy. Romania is a particularly relevant case because inflation has unfolded in parallel with persistent fiscal deficits, strong consumption-led growth, and repeated warnings from European and international institutions regarding macroeconomic imbalances and the need for fiscal consolidation.*

This article examines the impact of fiscal policy on inflation in Romania by combining theoretical discussion with recent macroeconomic evidence. It explains the main transmission mechanisms through which fiscal policy affects prices, analyses Romania's post-pandemic inflation dynamics, and evaluates the extent to which expansionary fiscal policy has contributed to inflation persistence. The paper argues that, although fiscal policy was not the only cause of inflation in Romania, it significantly amplified domestic inflationary pressures and reduced the speed and effectiveness of disinflation. The article also highlights that monetary policy alone cannot ensure price stability when the fiscal stance remains loose or inconsistent. Consequently, credible and well-designed fiscal consolidation is presented not only as a budgetary necessity, but also as an essential anti-inflation instrument.

Keywords: inflation, monetary policy, public deficit, sustainability, income, taxation.

JEL Classification: E31, E62, E63.

1. Introduction

Inflation is often treated as a technical concept, expressed through indices, percentage points, policy rates, and forecasts. Yet behind these figures lies a much more tangible reality: inflation shapes how households live, how businesses plan, how governments borrow, and how societies perceive economic stability. In a country like Romania—where living standards have improved significantly over the past two decades, but income vulnerability remains for many— inflation is not just a macroeconomic indicator. It is also a social and political concern. When prices rise too quickly, households lose purchasing power, businesses face cost uncertainty, and confidence in economic policy becomes harder to sustain.

In recent years, Romania has faced exactly this challenge. The post-pandemic recovery, the energy crisis, and the regional effects of the war in Ukraine all contributed to a surge in inflation across Europe. However, Romania's case stands out because inflation has remained closely tied to domestic vulnerabilities: high fiscal deficits, strong consumption, rapid increases in wages and pensions, and a broader imbalance between demand and the economy's productive capacity. Both the European Commission and the National Bank of Romania have repeatedly highlighted these imbalances, pointing to fiscal and income policies as key sources of inflationary pressure and uncertainty.

This context makes Romania a particularly relevant case for analyzing the relationship between fiscal policy and inflation. Fiscal policy is often discussed in terms of growth, redistribution, and public finance sustainability, but its role in shaping inflation is just as important. Government decisions on spending, taxation, wages, pensions, and borrowing directly influence demand, labor costs, inflation expectations, and the overall consistency of macroeconomic policy. In economies with persistent deficits and limited fiscal space, these effects become even stronger, as they interact with investor confidence, financing conditions, and exchange rate dynamics. Romania fits this profile closely.

The objective of this article is to examine the impact of fiscal policy on inflation in Romania, with a focus on recent developments while also drawing on economic theory and prior research. The central argument is that fiscal policy has not only accompanied inflation but has also contributed to its persistence—by sustaining domestic demand, reinforcing wage and income pressures, and weakening the overall policy mix aimed at price stability. The article also argues that lasting disinflation in Romania requires better alignment between fiscal and monetary policy, as central bank action alone cannot fully offset an expansionary or insufficiently credible fiscal stance.

The paper is structured as follows. The next section outlines the theoretical link between fiscal policy and inflation. This is followed by an analysis of Romania's recent inflation context and macroeconomic vulnerabilities. The subsequent sections examine the transmission channels of fiscal policy, assess the recent fiscal stance, and explore the interaction between fiscal and monetary policy. The paper then discusses the challenges of fiscal consolidation before concluding with the main findings and their policy implications.

2. Theoretical Framework and Romania's Inflation Context

Fiscal policy—how governments manage spending, taxation, and deficits—plays a central role in shaping inflation. At its core, the link is simple: fiscal decisions influence how much money flows into the economy, which in turn affects prices.

When governments increase spending, cut taxes, or expand social transfers, they boost purchasing power and stimulate demand. If this demand grows faster than the economy's ability to produce goods and services, prices rise. The opposite is also true: reducing spending or increasing taxes can ease inflationary pressure by slowing demand. This balance between demand and supply sits at the heart of how fiscal policy influences inflation (Blanchard, 2017).

In practice, inflationary effects become most visible when the economy is already under strain—when production cannot expand quickly, when imports are heavily relied upon, or when labor markets are tight. Under such conditions, additional demand translates more directly into higher prices rather than higher output. This dynamic is particularly relevant in Romania, where structural constraints and strong internal demand have amplified price pressures in recent years.

However, fiscal policy affects inflation not only through immediate demand, but also through expectations. When governments run large and persistent deficits, people begin to question whether such policies can be sustained. Businesses may anticipate future tax increases, workers may demand higher wages to protect their purchasing power, and investors may require higher returns to compensate for risk. These reactions feed into inflation itself, making it more persistent. In this sense, inflation reflects not only current economic conditions, but also confidence in future fiscal discipline.

Fiscal measures can also influence prices directly. Changes in indirect taxes, such as VAT or excise duties, are often passed on immediately to consumers. On the other hand, temporary interventions—such as price caps or subsidies—can delay inflation rather than eliminate it. Once these measures are removed, prices may adjust rapidly, sometimes creating new inflation spikes. This highlights how the design and timing of fiscal policy matter just as much as its size.

Another key element is the relationship between fiscal policy and monetary policy. In Romania, the National Bank uses interest rates to control inflation. But if fiscal policy continues to stimulate demand while the central bank is trying to slow it down, the two policies work against each other. In such cases, the central bank must act more aggressively, raising borrowing costs higher than would otherwise be necessary. This makes disinflation more difficult and potentially slows economic growth. Effective inflation control therefore depends on alignment between fiscal and monetary actions.

Romania's recent experience illustrates these mechanisms clearly. The country entered the post-pandemic period with existing vulnerabilities, including a large fiscal deficit, strong

reliance on domestic consumption, and a significant external imbalance. These weaknesses made the economy more sensitive to external shocks.

When global energy and food prices surged—especially after 2021—Romania experienced a sharp rise in inflation. But what followed is particularly important: inflation remained high even after external pressures began to ease. Headline inflation reached double-digit levels in 2022 and remained elevated in 2023, while core inflation—excluding volatile components like energy—stayed persistently high. This suggests that domestic factors were sustaining inflation beyond the initial shocks.

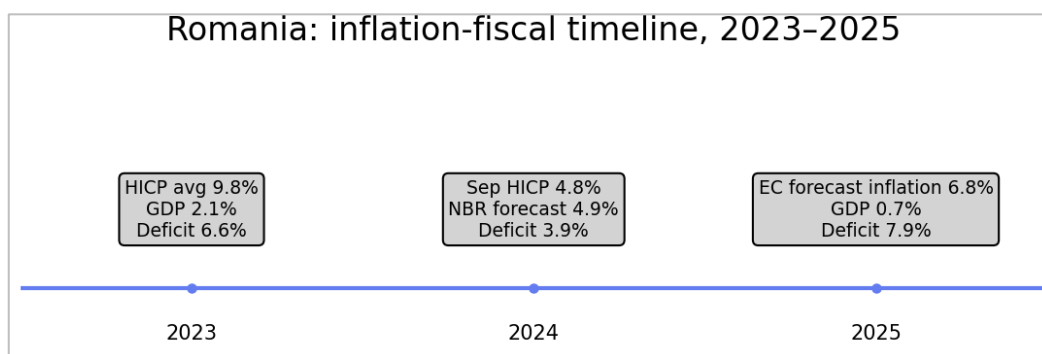
Several internal drivers contributed to this persistence. Wage growth accelerated, labor markets tightened, and both public and private consumption remained strong. At the same time, fiscal policy continued to support demand, while the budget deficit remained high and only gradually adjusted. This created a situation in which inflationary pressures were not just imported but increasingly generated within the economy itself.

Romania's position under the EU's Excessive Deficit Procedure further reflects these challenges. Persistent fiscal imbalances are not just a long-term sustainability issue—they also influence short-term inflation dynamics by shaping expectations and limiting policy flexibility.

Overall, Romania's inflation can be best understood as the result of overlapping forces. External shocks triggered the initial increase in prices. Domestic demand, supported in part by fiscal policy, sustained the pressure. And ongoing fiscal uncertainty reinforced inflation expectations. Together, these layers explain why inflation has been more persistent and harder to reduce than in a purely external shock scenario.

The key takeaway is clear: fiscal policy is not a passive background factor. It actively shapes inflation outcomes—either stabilizing the economy or making price control more difficult. In Romania's case, aligning fiscal policy with macroeconomic stabilization remains essential for achieving lasting price stability.

Figure 1. *Inflation and fiscal timeline in Romania, 2023–2025*



The figure combines historical observations with official forecasts to illustrate how disinflation, fiscal deterioration, and subsequent consolidation overlap. Note: the timeline mixes annual-average HICP (Harmonized Index of Consumer Prices), point-in-time HICP (Harmonized Index of Consumer Prices), and forecast values, so it should be presented as an interpretive macro timeline rather than a single-series statistical chart.

3. Transmission Channels: How Fiscal Policy Affected Inflation in Romania

3.1. Aggregate Demand and Consumption Pressure

The first and most obvious channel through which fiscal policy influences inflation is aggregate demand. When public expenditure rises rapidly, or when deficits remain large because spending exceeds revenue by a wide margin, the government effectively sustains demand in the economy. This effect becomes more pronounced if the additional spending directly increases household purchasing power or stimulates sectors already facing capacity constraints. In Romania, the economy has relied significantly on domestic consumption as a growth driver, and this feature makes inflation particularly sensitive to fiscal stimulus.

The European Commission's 2024 Country Report explicitly stated that growth in Romania was expected to be supported by stronger domestic demand, driven by higher real disposable income and public-sector related spending. This is a crucial observation, because it means that domestic demand was not naturally cooling enough to ensure a rapid return to price stability. Instead, fiscal and income policies contributed to preserving demand at relatively high levels. From an inflation perspective, this matters especially in an economy where supply-side rigidities and import dependence limit the speed with which higher demand can be met without price increases.

The OECD (Organization for Economic Cooperation and Development) position reinforces the same point. Its 2024 survey emphasized that fiscal consolidation would complement restrictive monetary policy in keeping demand at sustainable levels. This phrase is important because it acknowledges that inflation control depends not only on reducing the price of money through interest rates, but also on reducing the quantity of excess demand generated or sustained by fiscal policy. In Romania's case, the absence of sufficiently strong fiscal consolidation meant that monetary policy had to shoulder more of the anti-inflation burden.

3.2. Public Wages, Pensions, and Transfers

A second major channel operates through household income. In Romania, decisions regarding public wages, pensions, and other transfers have a broad macroeconomic impact because they affect disposable income for a large population segment. When such incomes increase quickly, consumption also tends to rise, especially in a context where households use a high share of income for current spending. If the supply side does not expand at the same pace, the imbalance pushes prices upward.

The European Commission explicitly warned that inflation could face upside risks if salaries and pensions continued to grow rapidly. This statement reflects the fact that wage and pension policy are not only social instruments; they are also macroeconomic variables with inflation consequences. Large increases in public-sector incomes can have two inflationary effects simultaneously: they directly increase household demand, and they influence wage-setting behavior in the wider economy, where private employers may face pressure to raise wages as well. In tight labor markets, such dynamics can feed broader cost pressures, especially in services.

The National Bank of Romania's own analysis supports this interpretation. Its Inflation Report included discussion of labor shortages and wage pressures as contributors to inflation. When the labor market is tight and productivity growth does not keep up with income growth, the result is a more inflationary internal environment. Fiscal policy becomes part of this mechanism whenever the public sector acts as a significant source of wage and transfer increases. In Romania, the interaction between fiscal income decisions and labor-market conditions has therefore been one of the main channels linking the budget to inflation.

3.3. Indirect Taxes, Excises, and Administered Price Effects

The third channel is more direct: fiscal policy can alter inflation through tax changes. Unlike aggregate demand effects, which work through broader macroeconomic transmission, tax changes affect prices almost immediately when businesses pass them on to consumers. VAT increases, excise adjustments, or changes in regulated pricing frameworks can produce a visible step-up in consumer prices. Conversely, temporary tax reductions or subsidies may dampen inflation for a period, even if the underlying pressures remain.

This distinction is especially important when discussing fiscal consolidation. In theory, deficit reduction should ease inflation in the medium term by reducing demand and improving credibility. In practice, however, consolidation that relies heavily on higher indirect taxes may increase inflation in the short term. The National Bank of Romania noted that fiscal and budgetary correction packages could alter the inflation outlook and that the future configuration of such measures remained a source of uncertainty. This means that fiscal adjustment can have mixed inflation effects depending on how it is designed.

The European Central Bank's material on Romania's price developments also helps clarify this issue by distinguishing between standard HICP (Harmonized Index of Consumer Prices) and HICP (Harmonized Index of Consumer Prices) at constant tax rates. This distinction highlights that changes in indirect taxation can create measurable shifts in consumer inflation independently of deeper demand conditions. For Romania, this means that policy evaluation must separate short-term price-level effects caused by tax measures from the broader medium-term role of fiscal policy in shaping aggregate demand and macroeconomic stability.

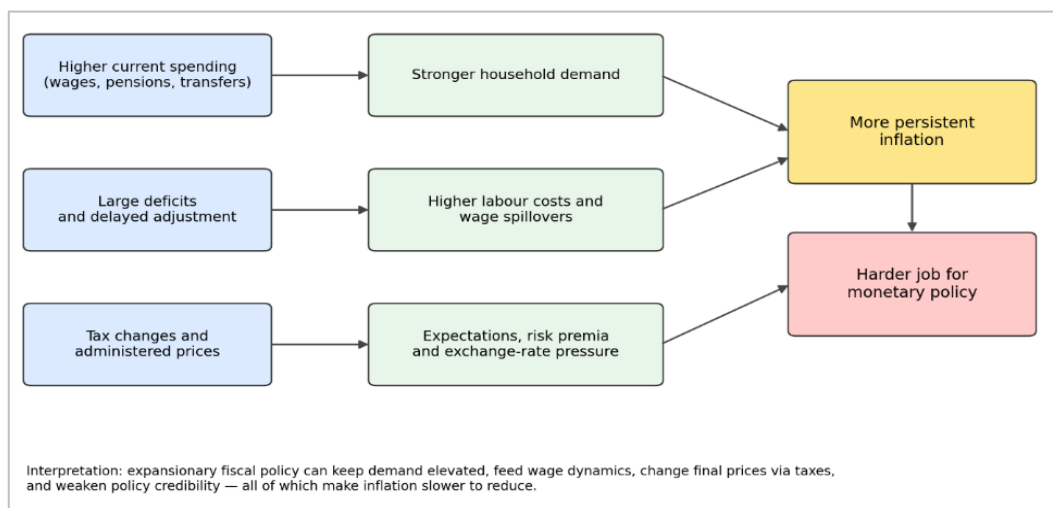
3.4. Expectations, Credibility, and External Vulnerability

A fourth and often underestimated channel concerns expectations and credibility. Inflation becomes more persistent when households, firms, and financial markets lose confidence that the policy mix is consistent. Persistent large deficits may signal that economic adjustment will be delayed, that public debt pressures will grow, or that taxes and administered prices may need to change abruptly later. Even if the government does not monetize the deficit, the perception of fiscal fragility can still influence price-setting, wage demands, and the exchange-rate environment.

Romania's continued position under the Excessive Deficit Procedure is relevant in this respect. The fact that the procedure remained open and that repeated steps were taken by the European institutions indicates not only a budgetary problem but also a credibility issue. If adjustment is repeatedly postponed, economic agents may begin to expect future corrections that are more abrupt and more painful. In emerging economies or converging economies such as Romania, this can influence inflation indirectly through financing conditions, exchange-rate sensitivity, and imported cost pressures.

The relevance of this channel is also supported by older analytical work on Romania. Budina and van Wijnbergen showed that fiscal deficits and inflation stabilization are interdependent, and that delayed adjustment raises the effort needed to restore consistency between fiscal and monetary objectives. Even though the institutional and European setting is different today, the underlying lesson remains valid: a credible fiscal anchor reduces inflationary risk, while a weak or delayed fiscal correction increases it.

Figure 2. Fiscal policy affects inflation in several ways in Romania



It boosts demand, puts pressure on wages, influences prices through taxes and regulated costs, and can reduce confidence in policy over time. Together, these effects make inflation slower to decline.

4. Romania's Recent Fiscal Stance and Its Inflationary Implications

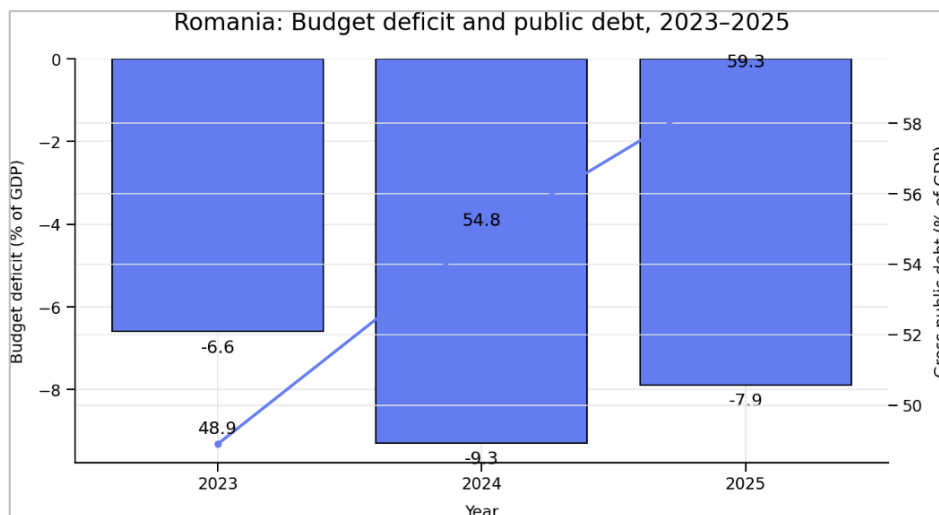
Romania's recent fiscal policy has often been seen by international institutions as not restrictive enough for the economic conditions it faced. This does not mean that the government intended to drive inflation higher. Rather, fiscal policy continued to support demand at a time when inflation control required a stronger correction. The OECD's 2024 assessment makes this point clear: tighter fiscal policy would have supported monetary tightening and helped keep demand at sustainable levels. The implication is straightforward—fiscal policy was still too expansionary for an economy dealing with high inflation.

The European Commission reached a similar conclusion. It linked strong domestic demand partly to public spending and to significant increases in wages and pensions. At the same time, it highlighted Romania's large current account deficit, which is closely connected to high government deficits. This matters because fiscal policy does not only affect the budget—it shapes the broader economy. By supporting consumption, it also increases reliance on imports, which adds pressure to external balances and makes inflation harder to bring down.

The persistence of Romania's excessive deficit reinforces this picture. European Commission reports show that the deficit has not been effectively corrected, despite repeated recommendations. This suggests that fiscal adjustment has lagged behind monetary policy. While the National Bank of Romania tightened monetary conditions to fight inflation, fiscal policy remained a source of demand pressure and uncertainty. This imbalance between policies helps explain why inflation has been slow to decline, even after passing its peak.

It is also important to separate the intent of fiscal policy from its effects. Many spending decisions—especially increases in wages, pensions, and social support—were driven by the need to protect households during a period of rising living costs. However, their macroeconomic impact depends on timing. When inflation is already high and demand is strong, such measures can unintentionally add to price pressures if they are not balanced elsewhere. This does not question the social value of these policies, but highlights their economic consequences in a given context.

Romania's experience reflects a broader policy dilemma. Governments aim to protect incomes during inflationary periods, but the tools used—higher wages, pensions, transfers, or tax relief—can also sustain demand and prolong inflation. The challenge is to strike a better balance: to support vulnerable groups without maintaining excess demand across the economy. International recommendations suggest that Romania has struggled to fully achieve this balance so far.

Figure 3. *Romania's budget deficit and public debt, 2023–2025*

The chart shows a sharp deterioration in the fiscal position in 2024, followed by a projected but still incomplete consolidation path in 2025–2025. Historical 2023–2025 figures are based on Eurostat-reported outcomes.

5. Fiscal Policy and Monetary policy: Complementarity or Conflict?

The relationship between fiscal policy and monetary policy is essential for understanding inflation in Romania. In countries with inflation targeting regimes, central banks are usually seen as the primary institutions responsible for price stability. This is also true in Romania, where the National Bank of Romania has operated under inflation targeting since 2005. However, central bank action can only be fully effective if it is not systematically undermined by fiscal expansion. A policy mix in which the central bank tightens while the government stimulates demand is coherent only for a limited time; beyond that point, the costs increase and the inflation outcome improves more slowly (Leeper, E. (1991) – “Equilibria under ‘Active’ and ‘Passive’ Monetary and Fiscal Policies”).

The NBR (National Bank of Romania) repeatedly emphasized that fiscal and budgetary policy was a major source of uncertainty for the inflation outlook. This is not a marginal technical comment; it reflects a structural point about macroeconomic coordination. When the central bank raises rates to control inflation, it seeks to reduce credit growth, moderate demand, and anchor expectations. If fiscal policy simultaneously raises household incomes or keeps the deficit too large, the central bank’s effort is partially offset. The result is that monetary tightening must remain in place for longer, which may affect investment, financing costs, and growth more heavily than if fiscal policy had adjusted earlier.

The OECD (Organization for Economic Cooperation and Development) expressed the same logic in more direct policy language. Its recommendation that fiscal consolidation

should complement restrictive monetary policy shows that the inflation problem was not viewed as purely monetary. Instead, price stability required a joint adjustment in which the government reduced the inflationary impulse of budget policy while the central bank maintained the nominal anchor. This is particularly relevant in Romania because the economy also faces current account pressures, meaning that excessive domestic demand has both internal and external consequences.

The interaction between the two policies is also important for credibility. When economic agents see a coherent macroeconomic mix, expectations stabilize more easily. When they see tension or contradiction between what the central bank is doing and what the government is doing, uncertainty increases. In Romania's case, the persistence of a high deficit while inflation remained elevated risked sending a mixed message: one institution was attempting to cool the economy, while the other was not adjusting fast enough. This tension may not always appear immediately in headline inflation figures, but it weakens the overall disinflation process.

A further point concerns the distribution of adjustment costs. If fiscal policy does not contribute enough to disinflation, the burden falls more heavily on interest rates, on the private sector, and on credit-sensitive activities. This means that households with loans, firms dependent on financing, and long-term investment projects carry more of the stabilization cost than they otherwise would. Better fiscal-monetary coordination is therefore not just a technical matter; it is also a way to distribute the burden of disinflation more efficiently and more fairly across the economy. Romania's recent experience suggests that a more balanced policy mix would have reduced some of these pressures.

6. Fiscal Consolidation in Romania: Necessity, Risks, and Design Principles

Given the preceding analysis, the argument for fiscal consolidation in Romania appears strong. A lower deficit should reduce aggregate demand pressure, improve fiscal credibility, support external rebalancing, and allow monetary policy to operate under better conditions. It should also strengthen Romania's position relative to European fiscal rules and reduce long-term debt risks. For these reasons, both the European Commission and the OECD (Organization for Economic Cooperation and Development) have argued in favor of credible budgetary adjustment.

Yet consolidation is not a simple mechanical solution. Much depends on how it is implemented. If the government relies primarily on indirect tax increases, the short-term effect on inflation may be adverse, even if the medium-term impact is stabilizing. If adjustment is based mainly on reducing productive public investment, long-term growth may suffer and supply capacity may weaken, which could create future inflationary pressures of another kind. If consolidation is abrupt and socially unbalanced, it may create political resistance strong enough to undermine credibility. Fiscal adjustment works best when it is not only large enough, but also well designed, gradual, and understandable to economic agents.

The OECD (Organization for Economic Cooperation and Development) provides a useful orientation regarding the quality of adjustment. It argues that Romania needs greater tax revenues to stabilize debt while continuing to fund priority spending such as infrastructure, pensions, health care, and education. This implies that the issue is not simply reducing spending across the board but improving the structure and efficiency of public finances. Stronger tax administration, broader revenue collection, a fairer tax base, and better control of non-productive expenditure can support consolidation without suffocating long-term development.

From an inflation perspective, the composition of fiscal consolidation matters enormously. Measures that reduce broad-based demand without directly increasing consumer prices are generally more supportive of disinflation than measures that immediately raise indirect taxes. Measures that protect public investment may also improve medium-term supply conditions, which helps weaken structural inflation pressure over time. Conversely, a poorly designed package may reduce the deficit but still leave the economy facing near-term price shocks, weaker growth, and fragile public support. Romania therefore needs not only consolidation, but smart consolidation.

The credibility dimension should not be underestimated either. Economic agents respond not only to enacted measures but also to whether they believe the government will maintain the adjustment path. Romania's prolonged deficit problem, already visible in the repeated stages of the Excessive Deficit Procedure, means that credibility has to be rebuilt, not assumed. A medium-term plan that is clear, realistic, and institutionally anchored would reduce uncertainty, improve coordination with monetary policy, and support lower inflation expectations. Without this credibility, even technically sound measures may have weaker effects than expected.

7. Is Fiscal Policy the Main Driver of Inflation in Romania? A Balanced Assessment

Although the evidence strongly suggests that fiscal policy has contributed to inflation persistence in Romania, the question must be answered with nuance. Fiscal policy was not the sole or original cause of the recent inflation episode. External shocks were clearly decisive in triggering the surge in prices. Energy markets, global supply disruptions, and regional geopolitical instability all had a major impact on European inflation, and Romania was no exception. The European Commission and the NBR (National Bank of Romania) both recognized the importance of these external influences when describing the inflation process.

However, fiscal policy should not be treated as peripheral simply because inflation had external starting points. What matters analytically is not only what started inflation, but also what allowed it to persist. In Romania, expansionary fiscal conditions, strong income growth, delayed deficit correction, and the lack of a fully coherent macroeconomic mix all

contributed to keeping domestic inflation pressure alive. This is particularly visible in the persistence of core inflation and in institutional repeated warnings about fiscal risks to the inflation outlook.

It is therefore more accurate to say that fiscal policy in Romania acted as an amplifier and prolonger of inflation rather than its sole origin. This distinction is important because it avoids exaggeration while still assigning responsibility where it exists. External shocks can ignite inflation, but domestic policy choices determine whether inflation fades naturally or becomes entrenched. In Romania's case, fiscal policy often tilted toward the latter outcome by supporting demand too strongly and by not adjusting deficits rapidly enough.

A balanced assessment also requires acknowledging that certain fiscal measures may have prevented social hardship in the short run. For example, income support and public expenditure can cushion the impact of inflation on vulnerable groups. But macroeconomic policy always involves trade-offs. Measures that protect income today may contribute to price pressure tomorrow if they are not targeted and financed credibly. This is exactly why the inflation-fiscal relationship is so complex and so important for Romania's policy future.

8. Broader Implications: Inflation as a Governance and Development Challenge

The Romanian case shows that inflation should not be viewed only as a short-term macroeconomic fluctuation. It is also a reflection of deeper issues in public finance governance, economic structure, and development strategy. An economy that relies heavily on consumption-led growth, persistent deficits, and recurrent policy improvisation will tend to be more vulnerable to inflation when shocks occur. By contrast, an economy with stronger productivity growth, better tax administration, disciplined current spending, and a credible policy framework will generally be more resilient. International institutions reviewing Romania's economy consistently emphasize these structural dimensions (European Commission, 2024).

This means that fighting inflation in Romania is not only about bringing monthly inflation rates down. It is also about improving the quality of governance. The issue includes how public money is spent, how revenues are collected, how medium-term budget commitments are made, how wage and pension policies are integrated into growth strategy, and how national policy aligns with European fiscal frameworks. The fact that Romania remained under the Excessive Deficit Procedure and continued to face repeated recommendations for adjustment indicates that the challenge is institutional as well as cyclical.

A second broader implication concerns convergence. Romania has made substantial progress in increasing GDP per capita and moving closer to European income levels, a point noted by the European Commission. Yet inflation can undermine that convergence if it erodes real incomes, damages competitiveness, and generates macro-financial vulnerability. In other words, inflation is not only a cost-of-living issue; it can become a drag on development itself if linked to structurally weak public finance management. Thus,

improving fiscal discipline and policy coherence is part of the broader challenge of ensuring that convergence remains sustainable.

Finally, Romania's experience offers a broader lesson for policy-making in emerging and converging economies. In such economies, the line between social support, growth support, and inflationary stimulus can be thin. The state must often respond to genuine social pressures, but those responses need to be designed in a way that does not undermine macroeconomic stability. This is why targeted support, credible medium-term planning, and transparent coordination between institutions are so important. Romania's inflation experience is, in this sense, a reminder that sound fiscal governance is a pillar of both stability and development.

9. Conclusion

The analysis in this paper shows that fiscal policy has played a significant role in shaping inflation in Romania, especially in the post-pandemic period marked by external shocks, strong domestic demand, and delayed fiscal adjustment. Fiscal policy influenced inflation through several channels: by sustaining demand, increasing disposable income through public wages and pensions, directly affecting prices through taxation, and shaping expectations through the credibility of the deficit path. These effects were particularly strong in Romania because they interacted with structural vulnerabilities such as current account imbalances, reliance on imports, and a consumption-driven growth model.

At the same time, the evidence does not support the idea that fiscal policy alone caused inflation. External shocks—especially energy and food price increases—were clearly responsible for the initial surge. However, fiscal policy amplified and prolonged these pressures by keeping demand elevated and weakening the overall policy mix. In this sense, fiscal policy was not the origin of inflation, but it became a key factor behind its persistence.

A central conclusion of this paper is that monetary policy alone cannot ensure a smooth and lasting disinflation if fiscal policy remains expansionary or uncertain. The National Bank of Romania can influence borrowing costs and anchor expectations, but it cannot fully offset strong fiscal support to demand. This is why international institutions have consistently emphasized the need for fiscal consolidation—not only to meet budgetary targets, but to support overall macroeconomic stability. Better coordination between fiscal and monetary policy would reduce the burden on the central bank, lower the cost of disinflation, and strengthen policy credibility.

At the same time, not all fiscal consolidation works equally well. The design of adjustment matters. Measures that improve tax collection, streamline spending, and protect productive investment are more effective than those relying heavily on higher indirect taxes or abrupt spending cuts. Romania therefore needs a fiscal strategy that is both corrective and forward-looking—one that reduces deficits while supporting long-term growth.

Ultimately, Romania's inflation experience shows that fiscal policy is not a secondary factor in price stability. It is one of its key drivers. While global shocks triggered inflation, domestic policy choices—especially around spending, deficits, and the timing of adjustment—shaped how long it lasted. Achieving lasting price stability will depend on turning fiscal policy from a source of pressure into a source of balance, credibility, and economic resilience.

References

- European Commission (2024). *2024 Country Report – Romania*. Brussels.
- European Commission (2024). *In-Depth Review 2024 – Romania*. European Economy Institutional Paper 275.
- European Commission. *Excessive Deficit Procedure and Romania*. Economy and Finance portal.
- National Bank of Romania (2024). *Inflation Report, November 2024*. Bucharest.
- OECD (Organization for Economic Cooperation and Development) (2024). *OECD (Organization for Economic Cooperation and Development) Economic Surveys: Romania 2024*. Paris: OECD (Organization for Economic Cooperation and Development) Publishing.
- Budina, N., and van Wijnbergen, S. (2000). *Fiscal Deficits, Monetary Reform, and Inflation Stabilization in Romania*. World Bank Policy Research Working Paper No. 2298.
- European Central Bank / Eurostat. *Romania – Price developments / HICP (Harmonized Index of Consumer Prices) context*.
- European Commission (2026 contextual country page, for convergence and later macro perspective). *Country report – Romania*.
- Leeper, E.M. (1991). Equilibria under “Active” and “Passive” Monetary and Fiscal Policies. *Journal of Monetary Economics*, 27(1), 129–147.

Comparing regional and global spillovers to the GCC equity markets

Moosa YOUSUF

University of Salford, United Kingdom
moosayo@outlook.com

Abstract. *This paper explores the global and regional financial connectedness for equity markets in the Gulf Cooperation Council (GCC) countries. Our empirical research benefits from two econometric approaches; the Diebold and Yilmaz (2012) connectedness approach and the DCC-GARCH model. These are applied to an extensive dataset of relevant MSCI indexes that consist of daily frequency data. For empirical analysis, the MSCI country level indexes are used to construct regional indexes separately for each country within the selected region. Our results indicate that the GCC equity markets are more strongly integrated intra-regionally, when compared against their level of integration with global markets.*

Keywords: Regional equity markets; international finance; equity market spillovers; financial connectedness; globalization.

JEL Classification: C55, C58, F65, G01.

1. Introduction

With globalization, financial markets around the world have become increasingly interconnected. In this context, a given equity market may be vulnerable to spillovers not only from within its geographic region, but also from much farther yet developed equity markets. Several regional studies have been conducted across equity markets representing different geographical cluster of countries. These include studies conducted by Mohammadi and Tan (2015), Chow (2017) and Yoon et al (2019) for a sample of Asian countries and by Ogbuabor et al (2016) for Africa. These studies show that geographical clusters could have a lesser influence on intra-regional connectedness than we would expect; certain equity markets within a particular geographic cluster might exhibit greater levels of connectedness with major global equity markets outside the cluster in contrast to other markets within the regional cluster, as illustrated by Sehgal et al (2018) for South Asian economies. Contrarily, Beirne et al (2010) illustrated that whilst global markets significantly influence the Asian equity markets, regional markets are the most significant determinants of stock market behaviour for other regions, such as the Latin American and Middle Eastern clusters. Therefore, existing literature demonstrates that the contrast between global and regional spillovers for an equity market is likely to significantly vary across different geographic regions.

This study focuses particularly on the six Gulf Cooperation Council (GCC) countries; UAE, Saudi Arabia, Qatar, Kuwait, Oman, and Bahrain, and examines how their equity markets respond differently to global and regional shocks. By consistently applying the same methodologies individually for each of the GCC countries' equity indexes over the same time horizon, we could contrast how these countries react differently (or similarly) to shocks from global and regional markets. To achieve this, we construct separate regional indexes for each of the six GCC countries, each excluding the corresponding country's index (for unbiasedness). A Tri-variate DCC GARCH Model is constructed for each of the six countries, including their respective regional indexes and the MSCI World Index that represents global equity markets in developed economies. This provides us with the dynamic conditional correlations that reveal how the local equity markets are integrated with regional and global markets over time. In addition, we apply the generalized VAR framework with rolling window estimations of forecast error variance decompositions, as suggested by Diebold and Yilmaz (2012), to evaluate the dynamic connectedness of local equity markets with regional and global markets. This enables exploration of the directional flow of spillovers for each of the GCC countries by constructing the net pairwise spillover plots. Finally, we compare the results obtained from these two completely different methods for robustness of our conclusions.

This research contributes to the literature by expanding our knowledge on the equity markets' behavior of GCC countries; we learn more about the time varying flow of regional and global equity market spillovers to the GCC economies. Whilst shocks from the developed equity markets could have a detrimental effect on emerging equity markets, this effect may surpass their intra-regional connectedness. However, we have reasons to hypothesize that this may not be the case for the GCC; high levels of economic integration and trade within the region, cultural and geopolitical similarities, and the shared impact

following oil price fluctuations. This research also offers insight on whether smaller equity markets are more prone to shocks from within the region or from global markets. Despite market size differences amongst the cluster of countries, they also vary in the degree of trade openness. Therefore, this study helps identify whether countries with higher trade openness are more susceptible to shocks from global markets. The time varying analysis of these spillovers also helps us identify how equity markets react during crises events. Ultimately, this study has the potential to help develop better informed macro-prudential policies that mitigate the effects of a financial contagion.

2. Literature Review

Since the global financial crisis of 2007 – 2008, many researchers in the fields of economics and finance have studied factors that could predict future crisis events and estimate the directional flow of spillovers in case of a financial contagion. Minoiu et al (2015) found that examining the connectedness of financial markets could help predict financial crises events. They found that the financial connectedness of markets sharply increased during the period preceding a systemic crisis event. Their research indicated that the financial connectedness of equity markets could serve as an early warning sign. Diebold and Yilmaz (2015) explored the interconnectedness between different financial markets using the connectedness approach as an indicator of financial integration and risk. Highly integrated markets are at greater risk of being impacted from a financial contagion in comparison to the relatively isolated markets. Moreover, whilst cross-border financial connectedness could be seen as a byproduct of globalization, financial markets have also become increasingly vulnerable to spillovers from other markets, therefore being subject to greater financial risks. To examine the financial integration levels of equity markets, researchers have applied various econometric techniques, of which, the GARCH models, and the connectedness approach by Diebold and Yilmaz, are two of the most popular ones.

One of the most widely used methods to examine spillovers across the various financial markets has been introduced by Diebold and Yilmaz (2009, 2010, 2012, 2015) and it will be later discussed in our methodology. Diebold and Yilmaz (2009) examined the case for 19 global equity markets and analyzed the connectedness of their returns and volatilities. They used vector auto-regression (VAR) models to extract forecast errors that could be divided into two parts; Errors explainable by shocks to other data series, and errors due to un-explained idiosyncratic factors. Through this formulation, it is possible to determine what proportion of shocks to a particular data series could be explained due to spillovers from others. For instance, Diebold and Yilmaz (2009) found that the US equity market had a high level of connectedness with other markets and was a net contributor of return and volatility shocks to other countries. Meanwhile, Japan had a relatively lower level of connectedness and was a net recipient of volatility shocks from other economies.

The Diebold and Yilmaz approach has also been adopted by many researchers who have examined the macroeconomic and financial connectedness of the financial markets across different countries and regions. Various researchers have also studied clusters of representative countries to examine the regional connectedness of their equity markets.

These studies include Guimaraes-Filho and Hong (2016) for East Asia, Fowowe and Shuaibu (2016) for Africa, Mensi et al (2018) and Skrinjaric and Sego (2019) for Europe, Demirer et al (2019) for the United States, Belcaid and El Ghini (2019) for Morocco, Chow (2017), Hsu and Lee (2019) and Yoon et al (2019) for Asia, and Schmidbauer et al (2016) for Russia. Whilst these researchers have specifically studied the equity markets, many authors have also examined the interconnectedness between equity markets and crude oil (Schmidbauer and Roesch, 2013; Maghyereh et al, 2016; Husain et al, 2019; Awartani and Maghyereh, 2013; Yousuf and Zhai, 2022).

Several authors have examined the interconnectedness of equity markets and other financial markets, such as the foreign exchange markets and bonds markets across a sample of countries. These include Abbas et al (2019) for the G7 countries, Sobti (2018) for India, Chau and Deesomsak (2014), Uluceviz and Yilmaz (2018) for the United States, Uluceviz and Yilmaz (2020) for Switzerland, Magkonis and Tsopanakis (2019) for Europe, and Fernández-Rodríguez and Sosvilla-Rivero (2020) for a global sample of seven countries. Some of the other authors also used the methodology to examine other areas such as real estate (Antonakakis et al, 2018; Jiang and Aijo, 2018), and business cycles (Ahmad and Sehgal, 2018). Abbas et al (2019) examined the dynamic spillover effects occurring to stock market returns and volatilities from six macroeconomic factors; industrial production, money supply, government bond rate, consumer price index, oil price and exchange rates. They conducted their empirical analysis individually for each of the G7 economies. Meanwhile, Yousuf (2024) examined the influence of news sentiments on the connectedness of GCC equity markets.

Current literature includes an important study on the GCC equity markets, conducted by Charfeddine and Al Refai (2019), who examined the connectedness between Qatar and the five other GCC countries between the years 2011 and 2018, in light of two specific crises events that happened during these times. The first of them referred to the unexpected large drop in oil prices during March 2014 and the second one referred to the June 2017 blockade on Qatar. Interestingly, during the first crisis in the time series, the total spillover connectedness within the region grew tremendously, whilst following the second crisis, it fell drastically. The authors examined the connectedness during three distinguished time periods: before the first crisis, between the first and second crisis, and after the second crisis. The level of connectedness between Qatar and other GCC countries did not change as much between the first two time periods as much as it did in comparison to the post second crisis event. One significant change following the first crisis was that the variance decomposition of Qatar stock returns showed an increase in directional connectedness from Dubai, from 6.7% to 16.5%. A major change after the blockade was that the directional connectedness from the GCC countries to Qatar significantly reduced, whilst remaining positive due to several reasons attributable to external sources, such as a common response to global oil price changes and the global investor sentiments about the overall geopolitical and economic risk within the region.

Another research on the GCC equity markets was conducted by Saadi-Sedik and Williams (2011) who apply GARCH models and find that the GCC equity markets are vulnerable to both intra-regional and extra-regional spillovers. They use the S&P 500 index as a proxy

for global spillovers, whilst other developed equity markets are excluded from the analysis. Their findings suggest that the inflow of shocks for the GCC markets peaked during the global financial crisis of 2007-2008. However, the time frame examined does not include major regional and global crises events that occurred over the past decade; the Arab Spring of early 2010s, the oil price crash of mid 2010s, the Qatar diplomatic crisis of 2017, and the global pandemic of 2020. This creates a gap in the literature that needs to explore the dynamic flow of connectedness and spillovers over the past decade. Furthermore, recent diversification strategies of GCC countries and an overall increase in the level of global financial integration may have caused significant changes to the connectedness dynamics and therefore, inferences from historic findings could be questionable. This provides us with an opportunity to explore the issues in further detail using a relatively recent, applying a combination of two different econometric methods.

In addition to Saadi-Sedik and Williams (2011), variations of the multivariate GARCH models have been popularly used by other researchers (Wang et al, 2005; Wang and Wang, 2010; Li and Giles, 2015). These models have become popular in modelling volatilities and forecasting due to their ease of application and the ability to control for serial autocorrelation of the variance term in an autoregressive moving average (ARMA) process. Furthermore, GARCH models have also been a preferred choice due to the variety of variations available such as the DCC-GARCH model that provides the dynamic conditional correlations, or the EGARCH model that captures asymmetric effects. The original GARCH model introduced by Bollerslev (1986), has been widely adapted due to its simplicity in application and the ability to control for heteroskedasticity. For these reasons, GARCH models have been widely adapted by researchers, especially for studies on stock market volatilities.

Bhar and Nikolova (2009) examined the dynamic correlation between return and volatility spillovers amongst BRIC countries using a bivariate EGARCH model. They found that amongst the BRIC countries, India had the strongest correlation with stock market volatilities, both within the cluster, and globally, followed by Brazil. China on the other hand exhibited the lowest level of integration and instead showed a negative correlation, suggesting that investors could consider this as an opportunity for diversification. EGARCH models are suitable towards computing volatility spillovers in stock markets, however further inferences regarding directional spillover effects and dynamic connectedness could be made based on the Diebold and Yilmaz approach. Similar diversification strategies have also been suggested by Mohammadi and Tan (2015) due to low correlation of equity market returns between the Chinese and US markets. Their conclusion was based on VAR models and three forms of GARCH models (BEKK, CCC and DCC). On the contrary, Hung (2019) used a Bi-variate BEKK-GARCH model and found strong unidirectional volatility spillovers from China to the Southeast Asian countries for their chosen time frame of study. A related study by Benlagha and El Omari (2021) conducted an analysis of five global stock markets using the DCC-GARCH approach. Their analysis also included two major commodities; oil and gold. Their results conformed to the intensification of spillovers during the recent crisis corresponding to the global pandemic, by intensification of the dynamic conditional correlations during the crisis event, compared to the pre-pandemic levels. They further used the connectedness

approach to conclude that the behaviour of both commodities differed from one another, such that oil was a net transmitter of shocks, whilst gold received shocks from the stock markets.

Researchers including Sehgal et al (2018) have utilized both, GARCH models and the Diebold and Yilmaz (2012) approach to examine the intra-regional connectedness of equity markets. They examine the time-varying connectedness amongst the equity markets of six South Asian countries; India, Pakistan, Bangladesh, Nepal, Sri Lanka, and Maldives. Unlike the results found for many other regions, the South Asian Association of Regional Cooperation (SAARC) member countries, exhibit a very low (near zero) level of financial integration. Sehgal et al (2018) have explained that the lack of intra-regional regional trade and investments amongst the SAARC countries could be a key explanatory factor for their results. Moreover, the author illustrates that the intra-regional connectedness amongst the equity markets of SAARC countries could also be significantly lower than the connectedness of SAARC countries with other Asian countries, such as member countries of the Association of Southeast Asian Nations (ASEAN). Both, GARCH models and the Diebold and Yilmaz approach, could be useful in providing a dynamic picture of the flow of spillovers. Assuming the connectedness is a dynamic concept that can rapidly change over time, we consider both these approaches suitable for our research. Yousuf and Zhai (2022) explored the global financial connectedness for GCC equity markets, utilizing both approaches. Doing so, they examined the connectedness of GCC equity markets with equity markets of the USA, Europe, and China, as well as crude oil. Extending their research, this paper includes separately constructed regional indexes for each of the six GCC countries to conduct a comparative study exploring global and intra-regional connectedness simultaneously.

The contrast between global and regional spillovers was also explored by Beirne et al (2010) who researched 41 emerging economies across four regions; Emerging Asia, Latin America, Emerging Europe, and MENA (Middle East & North Africa). They applied a trivariate GARCH model to examine the inflow of spillovers for each of the 41 emerging markets from the global equity market and from their respective regional equity market. They found mixed results for different countries, and in general, their results revealed that the Asian emerging markets were highly influenced by global spillovers, whilst the equity markets of Latin America and the Middle East, were relatively highly influenced by regional spillovers. Their sample included five countries from the Middle East; Egypt, Jordan, Kuwait, Lebanon and Saudi Arabia. Each of these countries, with the exception of Saudi Arabia, were significantly influenced by regional spillovers, whilst the Saudi equity market was influenced by global spillovers. Saudi Arabia being the largest equity market amongst the sample of Middle Eastern countries could be seen as an influential market causing intra-regional spillovers to other Middle Eastern equity markets, whilst itself being relatively highly influenced by the global equity market. Hence, smaller equity markets within the region may follow the key player resulting in directional spillovers from the Saudi equity market to other regional markets. A key emerging market within the region that was omitted in their study is of the UAE, which includes two major regional equity markets; the Dubai Financial Market (DFM) and the Abu Dhabi Securities Exchange (ADX).

Other studies that found evidence of spillovers from developed markets to emerging markets include Wang et al (2005) for equity markets of Pakistan, India and Sri Lanka, Hernandez et al (2020) for banks in South and North America, Li and Giles (2015) for China, India, Indonesia, Malaysia, Philippines and Thailand, and Yarovaya et al (2016) for a sample of 11 emerging markets across Asia, Europe, Africa and the Americas. For the banking sector, Hernandez et al (2020) examined the effect of spillovers from developed markets to emerging markets in the Americas and found that banks in developed markets had a strong influence on banks in emerging markets. Gębka and Serwa (2007) examined the intra-regional and inter-regional (global) connectedness for 12 emerging markets; four in Europe (Czech Republic, Hungary, Poland, Russia), four in Asia (Malaysia, South Korea, Taiwan, Thailand), and four in Latin America (Argentina, Brazil, Chile, Mexico). They found that whilst the correlation between equity markets of emerging economies can result from their common reaction to spillovers from the global capital markets, the spillovers from other regional markets can also be highly significant. They argue that equity markets sharing a common region are strategically interconnected due to similarities in their economic structures and policies. For these reasons, their results reveal that the intra-regional connectedness of equity markets could be significantly stronger than their inter-regional connectedness. Moreover, they extend their explanation to add that differences in trading hours due to having different time zones could be another significant factor contributing to higher intra-regional connectedness (and a relatively lower inter-regional connectedness). Overall, their results suggest that we could expect a higher level of intra-regional connectedness amongst the GCC equity markets, since the GCC countries not only operate under highly overlapping trading hours, but also have common economic structures and policies that distinguish the region from the rest of the world.

A related study for the GCC equity markets by Alotaibi and Mishra (2015) compared the impact of regional and global spillovers on the equity markets of five GCC countries; UAE, Qatar, Kuwait, Bahrain, and Oman. Their study utilizes the data for Saudi Arabia as a proxy for regional spillovers, since the Saudi equity market is considered as the largest equity market in the GCC, by market capitalization. Moreover, they consider the US equity markets as a proxy for global spillovers in their study. A major drawback in their study is the utilization of the Saudi equity market as a proxy for regional spillovers, omitting the study of global and regional spillovers for the Saudi equity market itself. This could be solved by introducing a separate GCC equity index for each of the GCC equity markets, by computing the market capitalization weighted average that excludes the equity market for which the index is being computed.

Our study on the global and regional spillovers is conducted using regional equity indexes established separately for each of the six GCC countries by computing the market capitalization weighted average for the other five markets. Moreover, we would utilize the MSCI World Index that covers a weighted average for the equity markets of 23 developed economies. Additionally, we apply both, the Diebold and Yilmaz (2012) approach, as well as the trivariate DCC-GARCH models to gather more robust and conclusive evidence on the impact of global and regional spillovers on the GCC equity markets. Our time varying analysis provides further insights on the dynamic flow of spillovers, illustrating how the inflow of regional and global spillovers to the GCC equity markets has evolved over time.

3. Data Collection and Description

Seven MSCI indexes are gathered for this study. Six of these are country-level indexes representing individual GCC member countries, and an additional index titled the MSCI World Index has been gathered as a representative of global equity markets (across 23 developed countries; excluding GCC countries). Our research utilizes these equity indexes for three reasons. First, the MSCI indexes would be gathered from a common source, computed using a standard approach, improving the comparability and reliability of our results. Second, the MSCI daily equity indexes were available for a longer timeframe, extending from 31st May 2005 until 31st December 2020. Third, the MSCI country-level index merges the equity indexes of DFM and ADX for the UAE, enabling a cross-country comparison.

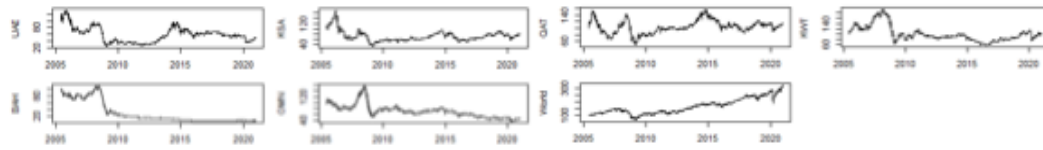
Our data covers stock market working days for a time frame extending over 15 years. This provides us with 4,069 observations (or 4,068 observations after taking the first difference). Hence, the daily-frequency data gathered provides sufficient observations to evaluate the time-varying effects of major global and regional events since mid-2000s. The MSCI equity indexes for each of the countries except for the UAE, represent their primary stock markets, whilst the UAE's index represents two stock markets: Dubai Financial Market (DFM) and the Abu Dhabi's Financial Market (ADX). Prior to further processing, data was indexed to set the base value of 100 on 31st May 2005, for each of the data series, by applying the equation $\bar{X}_t = \left(\frac{X_t}{X_0}\right) \times 100$, where $\bar{X}_t$ represents the indexed value for a series X_t . These indexes are plotted in Figure 1.

The MSCI World Index represents 23 developed markets, distributed across three regions classified by MSCI; Americas, Europe & Middle East, and Pacific. The MSCI World Index disproportionately represents each country by considering the market size, hence representing approximately 85% of the free float-adjusted market capitalization for each country. Therefore, the largest share of the index is represented by the US equity markets at 65.53%, followed by Japan, United Kingdom, France, and Switzerland at 7.96%, 4.41%, 3.41%, and 3.23% shares respectively. Hence, whilst the MSCI World Index is a representative index for the developed equity markets globally, nearly 85% of the index is computed based on the equity markets of these five countries alone. The distribution of representative countries for the MSCI World Index, is drawn in Figure 2. It is helpful to note that the representative equity markets for the MSCI World Index are primarily located in countries that have very different trading hours compared to the GCC, since researchers have argued that the time zone differences may also impact the connectedness of global equity markets, causing markets to be better integrated intra-regionally (Gębka and Serwa, 2007; Lyocsa et al, 2019).

Next, we utilize the six MSCI country-level equity indexes to construct market capitalization weighted regional GCC indexes that are constructed separately, for each of the six GCC countries. This is to avoid an individual country's index being included within its regional index resulting in autocorrelation. Hence, we construct separate regional indexes for each country, excluding the country for which, the index is constructed. Whilst doing so, it is important to note that the GCC stock markets vary greatly by market

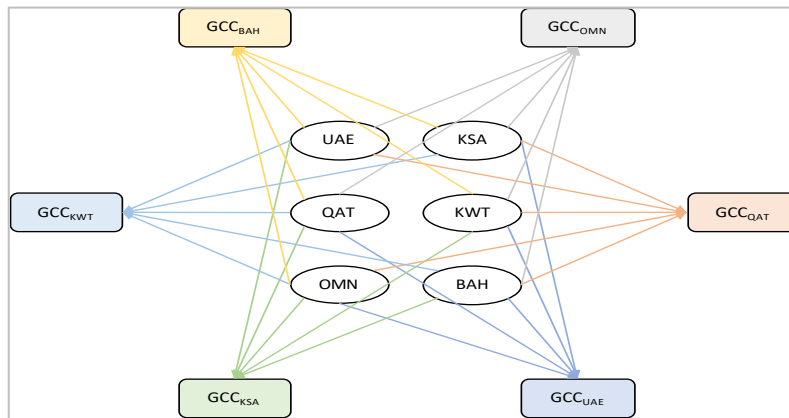
capitalization indicating significant differences in their market size. To account for these differences, regional indexes utilize weighted averages accounting for different market sizes. For example, the MSCI GCC Index is largely represented by the KSA's equity market due to its stock market being the largest regional market by market capitalization. Moreover, whilst the MSCI GCC Index exists, results from our connectedness study could be biased if this index were utilized due to autocorrelation. Hence, it is crucial to construct separate regional indexes in a manner outlined in Figure 2.

Figure 1. MSCI Country Indexes for the GCC and the MSCI World Index.



Note: These plots represent six country-level MSCI equity indexes, gathered for the six GCC countries, and the MSCI World Index, after indexing to set the base value of 100 for the first observation on 31st May 2005, for each series.

Figure 2. Illustration of the GCC Regional Indexes for Each GCC Country



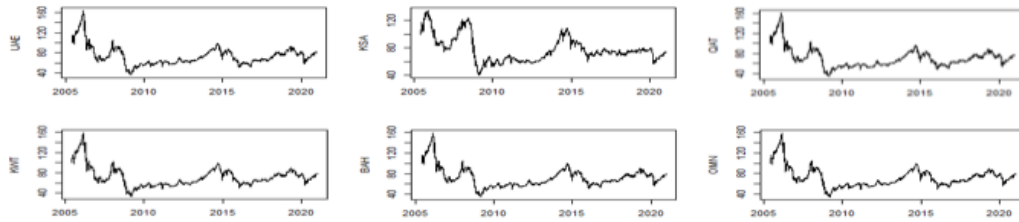
Note: This figure illustrates the use of five MSCI-country level indexes, in the construction of the GCC Regional Indexes for each of the six GCC countries. Hence, each regional index is constructed by taking the market capitalization weighted average for the other five GCC countries, to compute a regional index that excludes the country itself.

Our approach for the construction of regional indexes differs from other researchers such as Al Otaibi and Mishra (2015), who used the largest regional market as a proxy for the regional index, which in the case for the GCC is KSA. Such would prevent inclusion of the largest market, hence our approach effectively solves this issue. Therefore, to compute the regional index, we shall take the market capitalization weighted average of the “other five” countries for each of the six GCC countries. Weights would be attached based on market capitalization, to account for asymmetric market sizes within the region. Resultantly, we shall obtain six different regional indexes, for each country. The computation is based on the equation: $GCC Index_A = \sum_{i \neq A}^6 \left(\frac{Market\ Cap_i}{\sum_{i \neq A}^6 Market\ Cap_i} \times Country\ Index_i \right)$, where $GCC Index_A$ is the GCC Index

corresponding to country A . Once computed, the GCC regional indexes for each of the six countries, appear similar to each other, since each index includes a different set of five countries out of six, however, such an approach prevents us from the auto-correlation bias that would occur if the country being analyzed was also included in its regional index. The resultant indexes for each of the six GCC countries, are illustrated in Figure 3.

In order to extract the return on these indexes, we take the log deviation of each of the data series, to compute the daily rate of return for each of these indexes. This is done by simply computing $r_{it} = \ln(P_{it}) - \ln(P_{it-1})$, where P_{it} is the equity index for country or region i , at time t . The resultant daily equity return series for the MSCI country-level indexes are shown in Figure 4, alongside the resultant daily equity return series for the individual GCC equity indexes, as well as the MSCI World Index. Subsequently, the Dickey-Fuller Unit Root Test is applied on the return series, and stationarity is found, indicating that the return series are fit for time series analysis (Appendix A1). The descriptive statistics are provided in Table 1.

Figure 3. GCC Indexes constructed for each individual GCC countries separately



Note: The above plots represent GCC Indexes constructed separately for each country, using the market capitalization weighted average of the country-level equity indexes of all GCC countries, excluding the country for which the index is being generated (hence, representing all other regional equity markets, for each country).

Figure 4. Log Deviations of Daily Country Level, Regional, and MSCI World Indexes

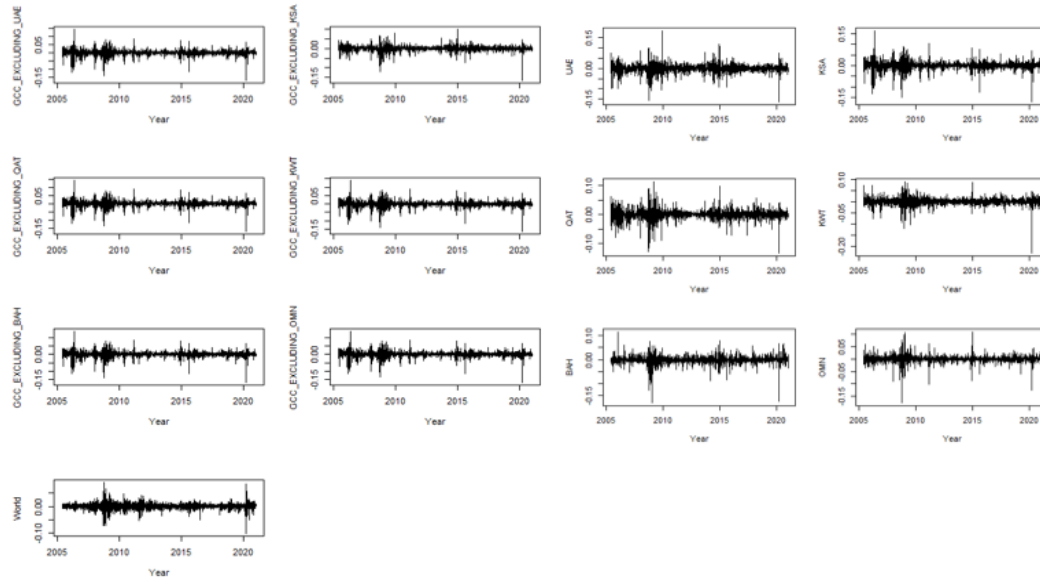


Table 1. Descriptive Statistics for the GCC country-level and regional indexes, as well as the MSCI World Index.

	MSCI KSA	MSCI UAE	MSCI QAT	MSCI KWT	MSCI BAH	MSCI OMN	
Mean	-0.0001	-0.0002	0.0000	0.0000	-0.0007	-0.0002	
Median	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	
Maximum	0.1623	0.1833	0.1127	0.0875	0.1150	0.1087	
Minimum	-0.1700	-0.1640	-0.1337	-0.2292	-0.1799	-0.1765	
Std. Dev.	0.0154	0.0168	0.0138	0.0130	0.0126	0.0114	
Skewness	-1.3961	-0.7324	-0.8609	-2.2499	-1.9637	-1.5842	
Kurtosis	20.5860	15.5296	14.9860	35.3967	31.0852	34.4132	
No. of Obs.	4,067	4,067	4,067	4,067	4,067	4,067	
	Regional Index KSA	Regional Index UAE	Regional Index QAT	Regional Index KWT	Regional Index BAH	Regional Index OMN	MSCI WORLD
Mean	-0.0001	-0.0001	-0.0001	-0.0001	-0.0001	-0.0001	0.0003
Median	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0007
Maximum	0.0999	0.1431	0.1433	0.1425	0.1368	0.1367	0.0910
Minimum	-0.1670	-0.1698	-0.1726	-0.1662	-0.1694	-0.1695	-0.1043
Std. Dev.	0.0114	0.0139	0.0140	0.0139	0.0135	0.0135	0.0106
Skewness	-1.8107	-1.6837	-1.6710	-1.6319	-1.7597	-1.7539	-0.7935
Kurtosis	25.5497	22.7738	22.4976	22.3104	23.3035	23.2423	13.4860
No. of Obs.	4,067	4,067	4,067	4,067	4,067	4,067	4,067

Note: The descriptive statistics are shown for each of the MSCI country-level indexes of each of the six GCC countries, as well as their market capitalization weighted regional equity indexes, and the MSCI World Index. The sample period ranges from 1st June 2005 until 31st December 2020.

4. Methodology

This research benefits from two different econometric methodologies that are applied to examine and compare the intra-regional and global spillovers for each of the six GCC countries. Each of these methods provide an opportunity to explore the intraregional connectedness of the regional equity markets and gain further insights. The first of these is proposed by Diebold and Yilmaz (2012), and the second is a Trivariate DCC-GARCH model proposed by Engle (2002). Both of these econometric methods allow us to explore the dynamic (time-varying) connectedness, whereby the tri-variate GARCH approach provides insights on the dynamic conditional correlations whilst the Diebold and Yilmaz (2012) approach provides further insights on the dynamic flow of spillovers as well as the net directional flow of these spillovers. These methods help us explore and compare between the magnitude of spillovers that the GCC equity markets receive from global markets versus other regional markets within the GCC. Both approaches are detailed in this section as follows.

5.1. The Diebold and Yilmaz (2012) Approach

Diebold and Yilmaz (2012) have used forecast error variance decompositions (FEVD) extracted from a Generalized Vector Autoregression (VAR) framework, introduced by Koop, Pesaran and Potter (1996) and Pesaran and Shin (1998). We obtain these forecast errors from the VAR models, in order to segregate the time series shocks that are due to spillovers from other variables in the system of equations. This methodology is detailed as follows:

Consider a simple VAR Model, given in Matrix Form, where X_t is an $(N \times 1)$ vector of dependent variables. The lagged values of X_t are represented by vectors from X_{t-1} until X_{t-p} , ε_t is an $(N \times 1)$ vector of error terms, p is the chosen lag length and γ_t is the matrix of coefficients at every lag:

$$X_t = \gamma_1 X_{t-1} + \dots + \gamma_p X_{t-p} + \varepsilon_t \quad (1)$$

In expanded matrix form, that would give:

$$\begin{bmatrix} X_{1t} \\ X_{2t} \\ \vdots \\ X_{Nt} \end{bmatrix} = \begin{bmatrix} \hat{\beta}_{1,1} & \dots & \hat{\beta}_{1,N} \\ \dots & \dots & \dots \\ \hat{\beta}_{N,1} & \dots & \hat{\beta}_{N,N} \end{bmatrix}_1 \begin{bmatrix} X_{1t-1} \\ X_{2t-1} \\ \vdots \\ X_{Nt-1} \end{bmatrix} + \dots + \begin{bmatrix} \hat{\beta}_{1,1} & \dots & \hat{\beta}_{1,N} \\ \dots & \dots & \dots \\ \hat{\beta}_{N,1} & \dots & \hat{\beta}_{N,N} \end{bmatrix}_p \begin{bmatrix} X_{1t-p} \\ X_{2t-p} \\ \vdots \\ X_{Nt-p} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \vdots \\ \varepsilon_{Nt} \end{bmatrix} \quad (2)$$

Next, consider a moving average representation for X_t that can be explained by observed values of the error terms:

$$X_t = \sum_i^{\infty} A_i \varepsilon_{t-i}; i > 0 \quad (3)$$

The distribution of coefficient matrix A_i follows the recursion, such that $A_i = \gamma_1 A_{i-1} + \gamma_2 A_{i-2} + \dots + \gamma_p A_{i-p}$. The moving averages component in the Generalized VAR framework, provides the coefficient matrices that can be used to extract forecast error variance decompositions, which are fundamental for the purpose of this study. An advantage of using the Generalized VAR framework to extract forecast error variance decompositions is that this framework is independent to the ordering of variables, unlike the previously applied Cholesky factorization (Diebold and Yilmaz, 2009). Finally, to compute the forecast variance decompositions, the forecast error variance $\varphi_{i,j}$ is estimated by choosing a specific forecast horizon H , whereby i represents the shock recipient and j represents the transmitter of shocks:

$$\varphi_{i,j}(H) = \frac{\sigma_{ii}^{-1} \sum_{h=0}^{H-1} (e_i' A_h e_j)^2}{\sum_{h=0}^{H-1} (e_i' A_h \sum A_h' e_i)} \quad (4)$$

In the above equation, σ_{ii} is the standard deviation of the error term for the i^{th} equation, e_i is a binary selection operator, which takes the value of 1 for the i^{th} element, or 0 otherwise. Furthermore, the summation takes the aggregate of different layers of the coefficient matrix A_h based on the selected forecast horizon. This is the variance matrix for the error term, that was initially computed as a part of the Generalized VAR framework. Therefore, the numerator is a representative of the total squared forecast errors occurring from j to i for a selected forecast horizon. These are scaled by the standard deviation and the denominator, which represents the total squared forecast errors occurring to i . However, based on the construct of $\varphi_{i,j}$, the row sums are not normalized yet. Hence, an extra step is taken to normalize each of the row sums to be equal to 1, so the values for $\varphi_{i,j}$ can make further interpretative sense, as a percentage figure.

$$a\tilde{\varphi}_{i,j} = \frac{\varphi_{i,j}}{\sum_{j=1}^N \varphi_{i,j}} \quad (5)$$

These forecast error variance shares can explain the extent to which, the errors occurring to a particular variable are caused by spillovers received from other variables in the system of equations. For all possible combinations of i and j , it is possible to construct an $N \times N$ table, which is referred to, as the connectedness table. The non-diagonal elements of this table represent directional spillovers from j to i , whilst each of the diagonal elements represent shocks to an individual series, which are not explainable by spillovers from others, but are instead, a result of various idiosyncratic factors. Hence, for a study involving five time series variables, the $N \times N$ matrix would appear as follows:

$$\begin{bmatrix} \tilde{\varphi}_{1,1} & \cdots & \tilde{\varphi}_{1,5} \\ \vdots & \ddots & \vdots \\ \tilde{\varphi}_{5,1} & \cdots & \tilde{\varphi}_{5,5} \end{bmatrix} \quad (6)$$

Whereby the non-diagonal row and column sums provide the net spillovers from and to other variables, respectively. For further clarification, these terms are explained as follows:

$$\text{Total spillovers from others for } i^{\text{th}} \text{ variable} = \sum_{\substack{j=1 \\ j \neq i}}^N \tilde{\varphi}_{i,j} \quad (7)$$

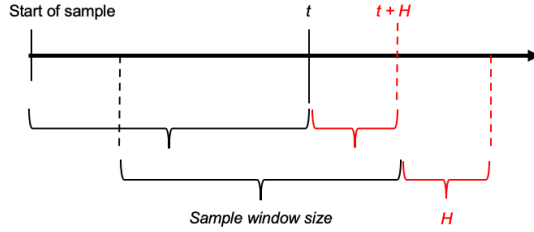
$$\text{Total spillovers to others for } j^{\text{th}} \text{ variable} = \sum_{\substack{i=1 \\ i \neq j}}^N \tilde{\varphi}_{i,j} \quad (8)$$

These values can be multiplied by a hundred, for ease of interpretation in percentages. The net spillovers for a particular variable is the difference between the above two, which denotes whether a particular financial market is a net recipient or a net transmitter of shocks. Furthermore, the total spillovers index can be computed by taking the summation of all non-diagonal elements of the connectedness matrix and dividing it by the summation of all elements of the matrix including the diagonals. By the construct of $\tilde{\varphi}_{i,j}$, as each row sums up to 1, all elements of the matrix sum up to N . Therefore, the total spillovers index is denoted as follows:

$$S = \frac{\sum_{i,j=1}^N (\tilde{\varphi}_{i,j})}{N} \times 100 \quad (9)$$

The total spillover index is a dynamic concept that is expected to vary over time and be affected due to major global and regional events. Therefore, the spillover index could be estimated repeatedly for rolling samples of a specific sample window length to examine the dynamic connectedness over an extended time frame. These sample windows can also be seen as a number of overlapping subsets of the overall time series data with equal number of observations. The rolling window method can be illustrated in Figure 5.

Figure 5. An illustration of rolling sample analysis



Note: This figure illustrates the rolling window approach for the estimation of forecast-error variance decompositions using rolling sub-samples of length labeled “sample window size”, which used to forecast a time horizon “H”, whilst the next sub-sample is selected for the same length, after every time interval equal to “H”.

The rolling sample method uses a specific sample window size, which in the above diagram could be denoted from the interval between the start of the sample until time “ t ”. This fixed interval is re-estimated with lags of “ H ”, which is the forecast horizon. For instance, data for an interval at time “ t ” is used to forecast for the following “ H ” number of stock market days. This is called an H -step ahead forecast, and is repeatedly estimated for equal sample window lengths, which are chosen at a specified lag “ H ”.

5.2. The DCC-GARCH Approach

The dynamic conditional correlation (DCC) GARCH models were introduced by Engle (2002), and have since, been widely adopted by researchers as a tool to measure time varying conditional correlations amongst stock market indexes. The DCC-GARCH model is a one of the many multivariate GARCH models, which are specified as follows:

$$r_t = \mu_t + \alpha_t \quad (10)$$

$$\alpha_t = H_t^{1/2} + \varepsilon_t \quad (11)$$

Where r_t is the vector of returns of N financial indexes at time t , μ_t is an $N \times 1$ vector of the expected value of the conditional r_t , whilst α_t is an $N \times 1$ vector of mean corrected return for r_t at time t ; such that the expected value of α_t is zero and its conditional variance is defined by the $N \times N$ matrix H_t . Note that $H_t^{1/2}$ can be obtained by a Cholesky factorization of H_t . Lastly, ε_t is an $N \times 1$ vector of independent and identically distributed errors.

There are various classes of multivariate GARCH models that differ in their specification of the conditional covariance matrix H_t . For our research, we will apply the DCC-GARCH model due its ability to model time varying conditional correlations, which would allow us to explore the dynamic flow of volatility spillovers. It is helpful to note that DCC-GARCH is an extension of the Constant Conditional Correlation (CCC) GARCH model, which specifies H_t as follows:

$$H_t = D_t R D_t \quad (12)$$

Whereby, D_t is the conditional standard deviation, represented by the diagonal elements of H_t , such that $D_t = \text{diag}(h_{1t}^{\frac{1}{2}}, h_{2t}^{\frac{1}{2}}, \dots, h_{Nt}^{\frac{1}{2}})$, whilst $R = [\rho_{ij}]$ is the positive definite constant correlation matrix with diagonal elements equating to 1, such that $\rho_{ij} = 1$ for every $i = j$. This ensures that H_t is always a positive definite matrix and its off-diagonal elements are given as:

$$[H_t]_{ij} = h_{it}^{\frac{1}{2}} \times h_{jt}^{\frac{1}{2}} \times \rho_{ij}; i \neq j \quad (13)$$

Although the CCC-GARCH model is easier and relatively straight-forward to compute, due to its correlation matrix R , being constant, it can be restrictive for the same reason. Therefore, the DCC-GARCH model allows for a time varying correlation matrix R_t , whereby R_t must be positive definite for t , for H_t to also be positive definite throughout the entire time series. Hence, the DCC-GARCH model is specified as follows:

$$H_t = D_t R_t D_t \quad (14)$$

Where R_t represents the conditional correlations matrix of standardized disturbances ϵ_t in α_t , such that $\epsilon_t = D_t^{-1} \alpha_t$ which has a normal distribution of $N(0, R_t)$. Therefore, R_t is a symmetric matrix, that also assumes the value of 1, for its diagonal elements:

$$R_t = \begin{bmatrix} 1 & \rho_{12,t} & \cdots & \cdots & \rho_{1N,t} \\ \rho_{12,t} & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \rho_{N-1,N,t} \\ \rho_{1N,t} & \cdots & \cdots & \rho_{N-1,N,t} & 1 \end{bmatrix} \quad (15)$$

Next, in order to ensure that R_t is a positive definite matrix that captures dynamic conditional correlations, R_t is defined by the DCC-GARCH model as a function of Q_t , which is the conditional covariance matrix of the standardized disturbances ϵ_t . Note that the conditional covariance matrix given below as Q_t , is time varying whilst the unconditional covariance matrix of ϵ_t given below as $\bar{Q}$, is constant over time.

$$R_t = \text{diag}(q_{11t}^{1/2} \dots q_{NNt}^{1/2}) * Q_t * \text{diag}(q_{11t}^{1/2} \dots q_{NNt}^{1/2}) \quad (16)$$

$$Q_t = (1 - a - b)\bar{Q} + a(\epsilon_{t-1} * \epsilon_{t-1}^T) + bQ_{t-1} \quad (17)$$

$$\bar{Q} = \text{COV}(\epsilon_t * \epsilon_t^T) = E(\epsilon_t * \epsilon_t^T) = \frac{\sum_{t=1}^T (\epsilon_t * \epsilon_t^T)}{T} \quad (18)$$

Note that introducing the diagonal matrix $\text{diag}(q_{11t}^{1/2} \dots q_{NNt}^{1/2})$ in R_t , provides a scaling effect that ensures that every element of R_t is either less than, or equal to 1. Furthermore, the scalar parameters a and b must be non-negative, whilst their sum must be less than 1, for the model to be stable. If a and b are both equal to zero, then by construction, we have $Q_t = \bar{Q}$, in which case, the CCC-GARCH model would be sufficient to estimate the correlation matrix H_t .

Whilst the DCC-GARCH model is well suited for modeling volatilities, a detailed account of dynamic return connectedness can be provided using the Diebold and Yilmaz (2012) approach. Using both approaches would allow us to examine the average connectedness between equity indexes over the time frame of our study, and the time varying flow of return spillovers across these indexes.

5. Empirical Findings and Discussion

Our research findings presented in this section contribute to the wider literature on financial connectedness, and in particular, extend the works of two prior studies that compared the global and regional spillovers for the GCC. Saadi-Sedik and Williams (2011) compared the global and regional spillovers for the GCC applying GARCH models on a dataset that extends until the end of 2000s, hence excluding major crises events that occurred over the past decade, such as the Arab Spring of early 2010s, the oil price crash of mid 2010s and the COVID-19 outbreak. Our study uses a larger and comparatively recent dataset, applying multiple methodologies. Meanwhile, Alotaibi and Mishra, (2015) utilized the largest regional equity market of Saudi Arabia as the proxy to examine regional spillovers, hence omitting the analysis of global and regional spillovers for the Saudi equity market itself. Our study includes each market through constructing a separate regional index for each country, including Saudi Arabia. Lastly, both studies assume the US equity markets as a sole representative of the global equity market, ignoring any spillovers that may arise from other global markets. Our study uses a rather more inclusive index; the MSCI World Index as a proxy for global spillovers. Using two distinct econometric approaches, we conduct the analysis separately for each of the 6 GCC countries. Results and corresponding discussion for each approach are presented independently in this section.

5.1. Results from the Diebold and Yilmaz (2012) Approach

In this analysis of global and regional spillovers for the GCC equity markets, we conduct the dynamic connectedness study, separately for each of the six GCC countries. The resultant connectedness tables are illustrated as Appendix A2, and the inflow of intra-regional and extra-regional spillovers, along with the net connected (computed as the difference between the outflow and inflow of spillovers), are summarized in Table 2.

Table 2. *Regional and Extra-regional Inflow of Spillovers and Connectedness*

	Inflow of regional spillovers	Net regional connectedness	Inflow of extra-regional spillovers	Net extra-regional connectedness
UAE	16.79	-0.09	7.61	-3.50
Saudi Arabia	17.81	-0.22	4.71	-0.84
Qatar	13.39	-0.54	7.07	-4.94
Kuwait	10.07	-0.50	2.28	0.83
Bahrain	6.33	-0.72	2.11	-0.11
Oman	11.45	-0.36	5.11	-3.40

Notes: The above table indicates the inflows and net inflows of spillovers from regional markets and global or extra-regional markets for each of the GCC's country-level indexes. The results represent a static overview of the connectedness for the 2005 – 2020 time period. Values for the construction of the above table are drawn from the individual connectedness tables for each country, constructed based on the Diebold and Yilmaz (2012) approach.

For each of the GCC countries, the inflow of regional spillovers are greater than the inflow of spillovers from equity markets in developed economies. This compliments our findings from the tri-variate GARCH models presented in the previous section. Moreover, the net regional connectedness for each of the countries is computed by deducting the outflow of spillovers from the inflow of spillovers. Each of the GCC countries are net recipients of spillovers, however the net regional connectedness is at low levels overall, suggesting that whilst these countries are recipients of shocks from within the region, they are also transmitters of shocks to other countries within the region. The lowest of these net regional connectedness figures corresponds to that of the UAE, at -0.09%. Being significantly close to zero, we could infer that whilst nearly 16.79 percent of shocks to the UAE could be explained by shocks occurring to other GCC countries, 16.70 percent of shocks occurring to other GCC countries can also be explained by shocks occurring to the UAE's equity markets, hence nullifying the net effect. On the other hand, the net inflow of spillovers from global markets is significantly higher, perhaps due to the absence of strong reciprocities of spillovers transmitted from the GCC to the rest of the world. For instance, whilst the UAE may transmit minor spillovers to other countries, the inflow of spillovers are 3.50 percent higher than the outflow of spillovers.

Overall, the UAE, Saudi Arabia and Qatar are the largest recipients of shocks from global markets, although the magnitude of these spillovers are smaller, compared to inflow of shocks from within the region, or also compared to other studies on other emerging markets that apply a similar methodology. Moreover, the UAE equity markets are the largest recipient of spillovers from the global markets as 7.61% of the forecast error variances for the UAE could be explained by fluctuations in the global markets. These findings complement the claims of Wang and Wang (2010), who support the argument that equity markets in countries with higher trade openness are relatively highly integrated with global markets. Following the UAE, the equity market of Qatar was the second highest recipient of spillovers from global markets, with 7.07% of the forecast error variances for Qatar, being explained by fluctuations in the global markets. It is insightful to note that the exact same ranking is also observed for the proportion of expats residing in the GCC, with the UAE taking the first place at 89%, followed by Qatar at 86% (Shayah and Sun, 2019).

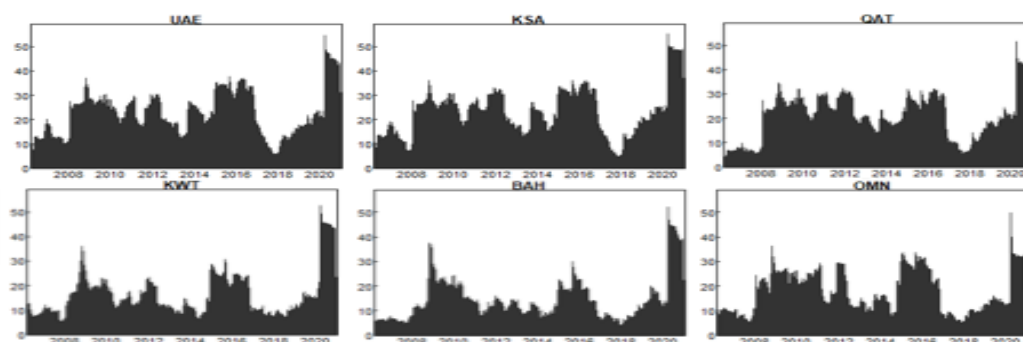
Interestingly, Bahrain received the least amount of spillovers from global markets at 2.11 percent, whilst it is also the largest net recipient of intra-regional shocks. Insightfully, it tells us a story. Since Bahrain is the smallest country within the region that forms an economic bloc, most of its trade occurs with Saudi Arabia; the largest country within the region and the only country it is geographically connected to, via land through a bridge. Much of the trade that occurs with global markets, also usually occurs via re-exports through other regional markets, especially Dammam, Saudi Arabia, and Dubai, United Arab Emirates. Hence, another hypothetical explanatory reason behind the low inflow of extra-regional spillovers to Bahrain, could be that these global spillovers may indirectly flow into Bahrain, getting accounted for, as regional spillovers. In either case, it makes intuitive sense that Bahrain may have higher net intra-regional spillovers and low extra-regional spillovers, compared to other GCC countries.

So far, we have evaluated the average connectedness throughout the entire time frame of our study. However, the connectedness levels can change over time. Hence, we also

examine the time varying connectedness using the rolling sample window approach, as elaborated in our methodology. The total dynamic connectedness plots (Figure 6), provide insights on the total level of connectedness across the three variables (domestic, regional and global indexes) for each of the six individual analyses. These plots reveal how the level of financial integration has varied over time. These can be interpreted as the proportion of forecast errors of the equity market returns that can be explained due to shocks occurring to other variables within the analysis. Subsequently, we can examine the specific net pairwise connectedness for the individual equity markets with the regional and global equity markets.

There are three major peaks across each of the total connectedness indexes for the six GCC countries, each representing one a global events; the global financial crisis of 2007-2008, the oil price crash of mid-2010s, and the global pandemic of 2020. These results complement the findings of other researchers including Diebold and Yilmaz (2015), who claimed that the connectedness of financial markets generally intensifies during crises events. Each connectedness index has peaked during the recent global pandemic, far beyond the previous crisis events, except for Oman. Ignoring a very slight peak of a few days, the impact of the global pandemic for Oman, is at levels comparable to the previous crises events. In addition to its low exposure to foreign markets compared to some of the other countries within the region, Oman also has a significantly low population density. Various studies have already shown a statistically significant and inverse correlation between population densities and the spread of COVID-19 (Kadi and Khelifaoui, 2020; Coşkun et al, 2020; Bray et al, 2020). This indicates more space and less crowdedness, which may have given Oman a comparative advantage in its fight against the virus. Unlike other crises events, the Qatar diplomatic crisis event showed a significantly noticeable decline in connected levels across all countries except for Kuwait and Oman. Although Kuwait and Oman also saw a decline in connectedness during this period, the magnitude of this decline was relatively lower. Intuitively, this would be expected since Kuwait and Oman were the only two GCC countries that did not participate in the blockade.

Figure 6. Total Connectedness Plots



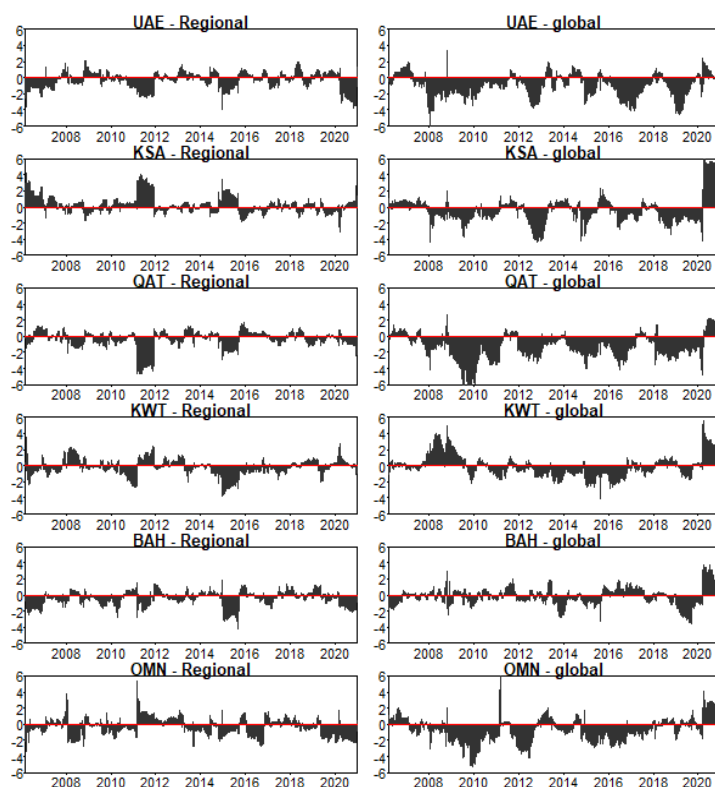
Note: The total connectedness plots represent the total dynamic connectedness for each GCC equity index, with their corresponding regional index and the MSCI World Index. The values could be interpreted as the total dynamic spillovers index, based on the Diebold and Yilmaz (2012) approach.

Overall, Saudi Arabia and the UAE exhibited the highest level of connectedness over the time frame, with an average total connectedness index of 18.99 percent and 18.28 percent,

respectively. On the other hand, Bahrain exhibited the lowest level of connectedness across all GCC countries, averaging at 9.21 percent, followed by Kuwait and Oman at 12.00 percent and 13.39 percent, respectively. In the study by Beirne et al (2010) that examined the inflow of global and regional spillovers for emerging markets including two countries from the GCC; Saudi Arabia and Kuwait, Saudi Arabia was the only one of the five Middle Eastern countries in the sample that received significant spillovers from global markets whilst the remaining Middle Eastern countries in their sample were more significantly influenced by regional spillovers. Whilst the UAE was not included in their study, it could also be expected for the UAE to top the chart for the total connectedness as Wang and Wang (2010) have argued that the higher connectedness could be associated with a greater degree of market openness, which is the highest for the UAE, amongst the GCC countries.

Whilst these total connectedness plots tell us about the general connectedness levels, without the global or regional distinction, we could obtain further insights from the net dynamic connectedness plots that our following discussion would be based on. The plots drawn in Figure 7, illustrate the time varying flow of net directional spillovers for between each country and the GCC regional index, as well as between each country and the MSCI World Index.

Figure 7. *Net Pairwise Connectedness Plots*



Note: The above plots represent the net dynamic pairwise connectedness for each GCC equity index, with their corresponding regional index (plots on the left) and the MSCI World Index (plots on the right). The amounts are computed based on the Diebold and Yilmaz (2012) approach, and could be interpreted in percentage terms. Values above zero, represent net transmission of shocks, and vice versa.

The net spillovers between most GCC countries and their respective regional indexes, represented by the plots on the left-hand side of Figure 7, appear to be closer to zero, compared to the net spillover plots for global markets on the right-hand side. This is even though the total inflow of spillovers from within the region, are higher, as shown from the connectedness tables. Perhaps, this could be explained by the bidirectional nature of spillovers within the region, in contrast to the predominantly unidirectional flow of spillovers from the global markets. During the aftermath of the global financial crisis of 2007-2008 and the Arab spring of early 2010s, Saudi Arabian equity markets take the role of net transmitter of spillovers, whilst the UAE and Qatar predominantly become the net recipients of shocks. The UAE and Qatar are both international markets that have significantly high proportions of foreign residents and foreign investment, not only compared to the rest of the GCC, but also in comparison with the rest of the world. Therefore, it makes intuitive sense that during times of political instability within the region, equity markets of these two countries may become more susceptible to spillovers from other markets. Furthermore, it also appears to be the case for the global inflow of spillovers that Qatar is the largest proportional net recipient of spillovers from global markets, followed by the UAE.

The net inflow of spillovers from global markets to the UAE, Qatar and Oman, were most prominent during the aftermath of the global financial crisis of 2007-2008, followed by the Arab spring of early 2010s and the oil price crash of mid-2010s. However, towards the end of the net global spillover plots for most of the GCC countries, there was an unexpected hike where the GCC appears to be a net transmitter of shocks to global markets, which is highly unlikely due to the relative market sizes. An alternative explanation for these hikes could be a common factor that raised the connectedness overall, which in this case was the COVID-19 outbreak. Since the COVID-19 outbreak had affected the Middle East earlier than the United States, the GCC equity markets could have captured this effect in the net-pairwise spillover plots. Here, it is helpful to note that the US Equity markets constitute of a significantly large proportion; 66.45% of the MSCI World Index and its COVID-19 infection rates for the year 2020, only peaked by mid-December, whilst the 2020 infection rates for Saudi Arabia, the largest GCC country by population, peaked 6 months earlier during mid-June. Therefore, the chronological order of the COVID-19 contagion across countries could explain the net-pairwise spillover plots for the year 2020.

5.2. Results from the Trivariate DCC-GARCH Models

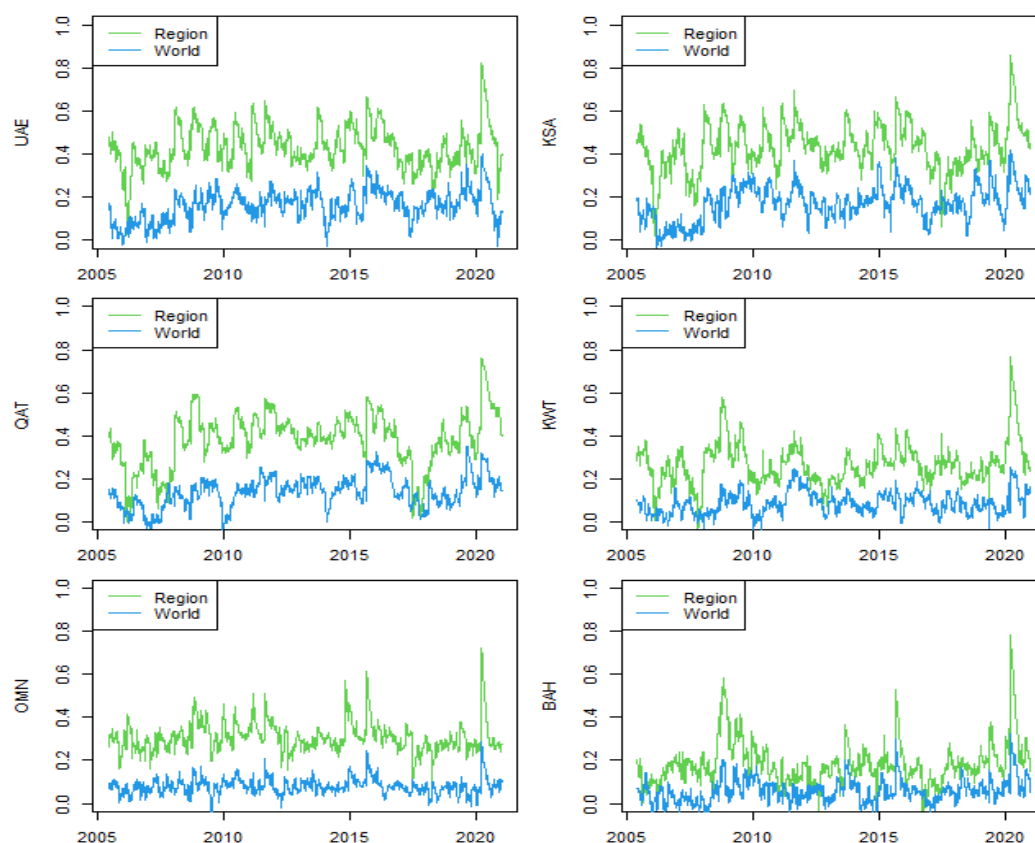
In the previous section, we explored the flow of spillovers for the GCC countries using the Diebold and Yilmaz (2012) approach. In this section, we can examine if our results from applying the tri-variate DCC-GARCH Models complement these findings. Each tri-variate DCC-GARCH analysis illustrates the regional and global intertemporal flow of dynamic conditional correlations, as demonstrated in Figure 10.

Several inferences could be made based on our outcomes of the tri-variate DCC GARCH models, presented in Figure 10. A general observation is that each country exhibits significantly higher intra-regional spillovers in contrast to spillovers from the global equity markets. This could potentially be due to the salient features of the GCC region, that not only functions as an economic bloc, but also shares common cultures, policy implications and geopolitical features. Furthermore, the GCC countries react differently to oil price

shocks, compared to most developed markets that use oil as inputs. Therefore, whilst it is expected that the countries within the GCC might be highly integrated within the region, the connectedness levels of these equity markets with the MSCI World Index may not be as high as the intra-regional connectedness.

Another observation made from our results is that the dynamic conditional correlations intensify during crises events. This compliments prior research by Yousuf and Zhai (2022), suggesting that regional equity markets become increasingly connectedness during times of uncertainty as they lack a strong sense of direction and look towards one another more so, than they otherwise would. This was most obvious for the global pandemic of 2020, the oil price crash of mid-2010s and the global financial crisis of 2007-2008. However, the global financial crisis of 2007-2008 did not have a prominent impact on the dynamic conditional correlations of Oman, possibly due to its relatively lower level of exposure to the equity markets in developed economies that represent the MSCI World Index.

Figure 10. *Dynamic Conditional Correlations with Regional and Global Markets*



Note: The above time-series plots illustrate the dynamic conditional correlations based on the multivariate DCC-GARCH models, computed for each individual GCC country. The regional time-series (in green), represents the dynamic conditional correlations for the given country's equity market with the equity markets of other GCC countries, whilst the "world" time-series (in blue), represents the dynamic conditional correlations for the given country, with the MSCI World Index.

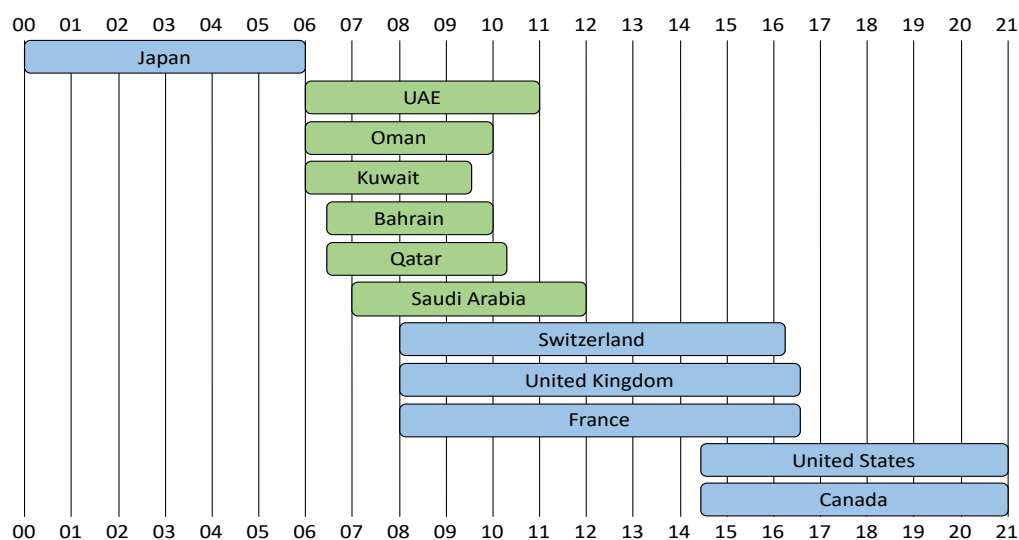
Impact of the COVID-19 outbreak had the highest effect on connectedness of GCC countries as the dynamic conditional correlations peaked for both global and regional equity markets. However, this impact was much stronger for the intra-regional spillovers as the dynamic conditional correlations reached an all-time high at values near 80% for all countries. On the other hand, the Qatar diplomatic crisis of 2017 had an inverse impact on the financial connectedness levels of GCC countries, especially for Qatar itself. This was the first time in a decade for Qatar that their dynamic spillovers from within the region plummeted to values near zero and dropped below the amount of dynamic spillovers from equity markets of developed economies from outside the region.

Interestingly, the two smallest equity markets within the region; Bahrain and Oman, exhibit the lowest levels of integration with the rest of the world. This contradicts the assumptions made by Wang and Wang (2010) that the high connectedness effect of emerging markets with the developed markets could be due to their small relative size. Our results show that in general, the GCC countries have stronger dynamic conditional correlations within the region, than external to the region. These overall results complement the findings of Saadi-Sedik and Williams (2011), and Alotaibi and Mishra (2015), who have found that the GCC equity markets are influenced by both regional and global spillovers. More specifically, our results support the claims of Beirne et al (2010) and Alotaibi and Mishra (2015), where they have argued that the inflow of spillovers from within the region, may be stronger than the inflow of spillovers from global markets external to the GCC. Additionally, our results also reveal that whilst the dynamic conditional correlations do not exhibit any strong patterns of growth for spillovers within the region, integration with global markets external to the region have trended upwards for the UAE, Saudi Arabia, and Qatar. Nonetheless, the spillovers from within the region remain high and above the spillovers from the MSCI World Index.

Our results complement the findings of Gębka and Serwa (2007) and Lyocsa et al (2019), who also argue that the intra-regional spillovers could be significantly stronger, whilst explaining that the time zone differences could account for the higher intra-regional connectedness amongst equity markets within the same region. Gębka and Serwa (2007) examined the case for equity markets across 12 emerging economies and claimed that whilst the equity markets of emerging economies can be correlated due to their common reaction to the developed equity markets, they are also likely to be strongly intra-regional connected due to common trading hours, in addition to having similar economic conditions and policies. To compare the trading hours of the GCC countries with the global markets, we first extract the trading hours of equity markets for each of the six countries and for the top six constituent countries that represent the MSCI World Index. Next, we convert each of their working hours into Greenwich Mean Time (GMT) for ease of comparison, as shown in Figure 10. Reviewing the trading hours on a common plot shows the discrepancies as pointed out by Gębka and Serwa (2007) for other emerging markets. The GCC countries not only share common economic structures and policies, but also operate during largely overlapping trading hours, as opposed to the markets represented by the MSCI World Index. Hence, our findings complement the claims of Gębka and Serwa (2007) and Lyocsa et al (2019), who argue that equity markets with largely overlapping

trading hours are also likely to be significantly interconnectedness, since the spillovers are expected to be higher for markets that operate during common trading hours.

Figure 10. *GCC Trading Hours versus the top six constituents of the MSCI World Index*



Note: All times are represented in GMT; the GCC countries are represented in green, whilst the top six constituents of the MSCI World Index are represented in blue; the trading hours for each country, are gathered from <https://www.tradinghours.com/markets>.

6. Conclusion

We have explored the interconnectedness for each of the GCC countries, against their respective GCC regional indexes and the MSCI World Index. Doing so, we implement tri-variate DCC-GARCH models and the Diebold and Yilmaz (2012) approach, both of which, are two of the most popular methods adapted by researchers to conduct similar studies. Our results from both methods, reveal that the GCC equity markets to varying degrees, are affected by shocks from both regional and global markets. Furthermore, the amount of spillovers from regional markets are higher than the extra-regional spillovers, perhaps due to the distinguishing features of the GCC countries that they share. Therefore, the GCC markets continue to provide diversification opportunities to international investors and portfolio managers. However, exploring individual countries and sectors within the GCC would be essential, since despite common trends, there are several idiosyncrasies that also prevail. For instance, whilst the equity markets of Bahrain have the lowest extra-regional inflow of spillovers compared to the other GCC countries, Bahrain is also the highest net recipient of regional shocks. Meanwhile, the UAE and Qatar are the two largest net recipients of shocks from the global markets, which makes intuitive sense, due to their exposure to foreign markets and high international trade openness levels.

As countries became more and more isolated due to the recent outbreak of the COVID-19, their financial markets have become more connected than they have ever been in over a

decade. Both stock market returns and volatilities of the GCC countries closely follow each other during crises events more than they would otherwise normally do so. Our results complement Charfeddine and Al Refai (2019) and Yousuf and Zhai (2022) regarding the intensification of spillovers during significant global and regional events, including the inverse effect of the Qatar Diplomatic crisis. Interestingly, all countries except for Oman, show that the impact of the recent global pandemic was significantly higher than any of the previous crisis events over the entire time frame of our study. As for Oman, low population densities and low exposure to global markets compared to most of the other GCC countries, could've have acted as a semi-shield against a larger contagion. Overall, the total connectedness index has exhibited an upward trend over the past few years. With the continuously increasing levels of financial integration, entirely avoiding a contagion corresponding to a financial crisis may no longer be an option. Nonetheless, examining the directional flow of spillovers will help foresee the transmission of spillovers if an exogenous shock occurs to a specific equity market or region.

References

- Abbas, G. et al. (2019). Return and Volatility Connectedness between Stock Markets and Macroeconomic Factors in the G-7 Countries. *Journal of Systems Science and Systems Engineering*, 28(1), 1–36.
- Ahmad, W., and Sehgal, S. (2018). Business Cycle and Financial Cycle Interdependence and the Rising Role of China in SAARC. *Journal of Quantitative Economics*, 16(2), 337–362.
- Alotaibi, A.R., and Mishra, A.V. (2015). Global and regional volatility spillovers to GCC stock markets. *Economic Modelling*, 45, 38–49.
- Antonakakis, N., Chatziantoniou, I., Floros, C., and Gabauer, D. (2018). The dynamic connectedness of UK regional property returns. *Urban Studies*, 55(14), 3110–3134.
- Awartani, B., and Maghyereh, A.I. (2013). Dynamic spillovers between oil and stock markets in the Gulf Cooperation Council Countries. *Energy Economics*, 36, 28–42.
- Belcaid, K., and El Ghini, A. (2019). Spillover Effects among European, the US and Moroccan Stock Markets before and after the Global Financial Crisis. *Journal of African Business*, 20(4), 525–548.
- Benlagha, N., and El Omari, S. (2021). Connectedness of stock markets with gold and oil: New evidence from COVID-19 pandemic. *Finance Research Letters*, 102373.
- Beirne, J., Caporale, G.M., Schulze-Ghattas, M., and Spagnolo, N. (2010). Global and regional spillovers in emerging stock markets: A multivariate GARCH-in-mean analysis. *Emerging markets review*, 11(3), 250–260.
- Bhar, R., and Nikolova, B. (2009). Return, volatility spillovers and dynamic correlation in the BRIC equity markets: An analysis using a bivariate EGARCH framework. *Global Finance Journal*, 19(3), 203–218.
- Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics*, 31(3), 307–327.

- Bray, I., Gibson, A., and White, J. (2020). Coronavirus disease 2019 mortality: a multivariate ecological analysis in relation to ethnicity, population density, obesity, deprivation and pollution. *Public health*, 185, 261-263.
- Charfeddine, L., and Al Refai, H. (2019). Political tensions, stock market dependence and volatility spillover: Evidence from the recent intra-GCC crises. *The North American Journal of Economics and Finance*, 50, 101032.
- Chau, F., and Deesomsak, R. (2014). Does linkage fuel the fire? The transmission of financial stress across the markets. *International Review of Financial Analysis*, 36, 57-70.
- Chow, H.K., (2017). Volatility spillovers and linkages in Asian stock markets. *Emerging Markets Finance and Trade*, 53(12), 2770-2781.
- Coşkun, H., Yıldırım, N., and Gündüz, S. (2021). The spread of COVID-19 virus through population density and wind in Turkey cities. *Science of the Total Environment*, 751, 141663.
- Demirer, M., Gokcen, U., and Yilmaz, K. (2019). Financial sector volatility connectedness and equity returns. Available at SSRN 3099738.
- Diebold, F.X., and Yilmaz, K. (2009). Measuring financial asset return and volatility spillovers, with application to global equity markets. *The Economic Journal*, 119(534), 158-171.
- Diebold, F.X., and Yilmaz, K. (2012). Better to give than to receive: Predictive directional measurement of volatility spillovers. *International Journal of forecasting*, 28(1), 57-66.
- Diebold, F.X., and Yilmaz, K. (2015). Financial and macroeconomic connectedness: A network approach to measurement and monitoring. *Oxford University Press*, USA.
- Engle, R. (2002). Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business & Economic Statistics*, 20(3), 339-350.
- Fernández-Rodríguez, F., and Sosvilla-Rivero, S. (2020). Volatility transmission between stock and foreign exchange markets: a connectedness analysis. *Applied Economics*, 52(19), 2096-2108.
- Fowowe, B., and Shuaibu, M. (2016). Dynamic spillovers between Nigerian, South African and international equity markets. *International Economics*, 148, 59-80.
- Gębka, B., and Serwa, D. (2007). Intra-and inter-regional spillovers between emerging capital markets around the world. *Research in International Business and Finance*, 21(2), 203-221.
- Guimaraes-Filho, R., and Hong, G.H. (2016). Dynamic Connectedness of Asian Equity Markets. IMF Working Papers, IMF Working Papers, March 10, 2016.
- Hernandez, J.A., Kang, S.H., Shahzad, S.J.H., and Yoon, S.M. (2020). Spillovers and diversification potential of bank equity returns from developed and emerging America. *The North American Journal of Economics and Finance*, 54, 101219.
- Hsu, S.-H., and Lee, J.-C. (2019). Measuring the Connectedness of Returns and Volatilities in Asian Stock Markets. *Jing Ji Lun Wen Cong Kan*, 47(4), 579-620.
- Hung, N.T. (2019). Return and volatility spillover across equity markets between China and Southeast Asian countries. *Journal of Economics, Finance and Administrative Science*.
- Husain, S., Tiwari, A.K., Sohag, K., and Shahbaz, M. (2019). Connectedness among crude oil prices, stock index and metal prices: An application of network approach in the USA. *Resources Policy*, 62, 57-65.
- Jiang, J., and Aijo, J. (2018). Equity volatility connectedness across China's real estate firms and financial institutions. *Journal of Chinese Economic and Business Studies*, 16(2), 215-231.
- Kadi, N., and Khelfaoui, M. (2020). Population density, a factor in the spread of COVID-19 in Algeria: statistic study. *Bulletin of the National Research Centre*, 44(1), 1-7.

- Koop, G., Pesaran, M.H., and Potter, S.M. (1996). Impulse response analysis in nonlinear multivariate models. *Journal of econometrics*, 74(1), 119-147.
- Li, Y., and Giles, D.E. (2015). Modelling volatility spillover effects between developed stock markets and Asian emerging stock markets. *International Journal of Finance & Economics*, 20(2), 155-177.
- Lyocsa, S., Vyrost, T., and Baumohl, E. (2019). Return spillovers around the globe: A network approach. *Economic Modelling*, 77, 133-146.
- Maghyereh, A.I., Awartani, B., and Bouri, E. (2016). The directional volatility connectedness between crude oil and equity markets: New evidence from implied volatility indexes. *Energy Economics*, 57, 78-93.
- Magkonis, G., and Tsopanakis, A. (2019). The financial connectedness between Eurozone core and periphery: A disaggregated view. *Macroeconomic Dynamics*, 1-26.
- Mensi, W., Boubaker, F.Z., Al-Yahyaee, K.H., and Kang, S.H. (2018). Dynamic volatility spillovers and connectedness between global, regional, and GIPSI stock markets. *Finance Research Letters*, 25, 230-238.
- Minoiu, C., Kang, C., Subrahmanian, V.S., and Berea, A. (2015). Does financial connectedness predict crises? *Quantitative Finance*, 15(4), 607-624.
- Mohammadi, H., and Tan, Y. (2015). Return and volatility spillovers across equity markets in mainland China, Hong Kong and the United States. *Econometrics*, 3(2), 215-232.
- Ogbuabor, J.E., Orji, A., Aneke, G.C., and Erdene-Urnukh, O. (2016). Measuring the real and financial connectedness of selected African economies with the global economy. *South African Journal of Economics*, 84(3), 364-399.
- Pesaran, H.H., and Shin, Y. (1998). Generalized impulse response analysis in linear multivariate models. *Economics letters*, 58(1), 17-29.
- Saadi-Sedik, T., and Williams, O.H. (2011). Global and regional spillovers to GCC equity markets. *IMF Working Papers*, 1-26.
- Schmidbauer, H., and Roesch, A. (2013). International connectedness of equity markets and the crude oil market: the case of Iran, Pakistan, and Turkey. IDEAS Working Paper Series from RePEc, IDEAS Working Paper Series from RePEc, 2013.
- Schmidbauer, H., Röscher, A., Uluceviz, E., and Erkol, N. (2016). The Russian stock market during the Ukrainian crisis: A network perspective. *Finance a Uver*, 66(6), 478.
- Sehgal, S., Pandey, P., and Deisting, F. (2018). Time varying integration amongst the South Asian equity markets: An empirical study. *Cogent Economics & Finance*, 6(1), 1-26.
- Shayah, M.H., and Sun, Z. (2019, January). Employment in the Gulf Cooperation Council (GCC) Countries—Current Issues and Future Trends. In *2nd International Conference on Social Science, Public Health and Education (SSPHE 2018)* (412-415). Atlantis Press.
- Sobti, N. (2018). Domestic intermarket linkages: measuring dynamic return and volatility connectedness among Indian financial markets. *Decision*, 45(4), 325-344.
- Uluceviz, E., and Yilmaz, K. (2018). Measuring Real-Financial Connectedness in the U.S. Economy. IDEAS Working Paper Series from RePEc, IDEAS Working Paper Series from RePEc, 2018.
- Uluceviz, E., and Yilmaz, K. (2020). Real-financial connectedness in the Swiss economy. *Swiss Journal of Economics and Statistics*, 156(1), 1-20.
- Wang, P., and Wang, P. (2010). Price and volatility spillovers between the Greater China Markets and the developed markets of US and Japan. *Global Finance Journal*, 21(3), 304-317.

- Wang, Y., Gunasekarage, A., and Power, D.M. (2005). Return and volatility spillovers from developed to emerging capital markets: the case of South Asia. In *Asia Pacific Financial Markets in Comparative Perspective: Issues and Implications for the 21st Century*. Emerald Group Publishing Limited.
- Yarovaya, L., Brzeszczyński, J., and Lau, C.K.M. (2016). Intra-and inter-regional return and volatility spillovers across emerging and developed markets: Evidence from stock indices and stock index futures. *International Review of Financial Analysis*, 43, 96-114.
- Yoon, S.M., Al Mamun, M., Uddin, G.S., and Kang, S.H. (2019). Network connectedness and net spillover between financial and commodity markets. *The North American Journal of Economics and Finance*, 48, 801-818.
- Yousuf, M., and Zhai, J. (2022). The financial interconnectedness between global equity markets and crude oil: evidence from the GCC. *Journal of Chinese Economic and Business Studies*, 1-24.
- Yousuf, M. (2024). The Role of News Sentiments in the Connectedness of GCC Equity Markets. *Journal of Economic Analysis*, 3(4), 136-159.

Appendixes

A1. Unit Root Testing.

	t-Stat values for variables used in the daily dataset			
	Observations = 4,066 no constant, no lag	Observations = 4,065 no constant, with lag	Observations = 4,066 with constant, no lag	Observations = 4,065 with constant, with lag
UAE	-59.031	-41.039	-59.03	-41.04
KSA	-62.072	-43.325	-62.066	-43.321
QAT	-60.434	-42.343	-60.427	-42.338
KWT	-61.704	-42.995	-61.697	-42.989
BAH	-61.662	-42.953	-61.832	-43.127
OMN	-58.485	-41.226	-58.492	-41.234
GCC_EXCL_UAE	-61.588	-42.825	-61.581	-42.821
GCC_EXCL_KSA	-58.991	-39.885	-58.986	-39.883
GCC_EXCL_QAT	-61.526	-42.822	-61.52	-42.818
GCC_EXCL_KWT	-61.184	-42.536	-61.178	-42.533
GCC_EXCL_BAH	-61.163	-42.465	-61.156	-42.461
GCC_EXCL_OMN	-61.181	-42.482	-61.175	-42.479
WORLD	-60.178	-43.421	-60.212	-43.459
Critical Values for Dickey-Fuller t-Dist				
	1%	5%	10%	
No constant	-2.58	-1.95	-1.62	
With constant	-3.43	-2.86	-2.57	

A2. Tables for Global and Intra-regional Connectedness.

The following results illustrate six connectedness tables representing the global and intra-regional connectedness for the six MSCI country-level indexes of the GCC countries.

		UAE	GCC	WORLD	FROM OTHERS
UAE	UAE	75.59	16.79	7.61	24.41
	GCC	16.7	77.98	5.32	22.02
	WORLD	4.11	4.31	91.58	8.42
	TO OTHERS	20.82	21.10	12.93	TCI
	NET DIRECTIONAL	-3.59	-0.92	4.51	18.28
KSA	KSA	77.48	17.81	4.71	22.52
	GCC	17.59	73.86	8.55	26.14
	WORLD	3.87	4.44	91.69	8.31
	TO OTHERS	21.46	22.25	13.26	TCI
	NET DIRECTIONAL	-1.06	-3.89	4.95	18.99
QAT	QAT	79.54	13.39	7.07	20.46
	GCC	12.85	81.30	5.85	18.70
	WORLD	2.13	4.59	93.28	6.72
	TO OTHERS	14.98	17.99	12.92	TCI
	NET DIRECTIONAL	-5.48	-0.72	6.2	15.30
KWT	KWT	87.65	10.07	2.28	12.35
	GCC	9.57	83.95	6.48	16.05
	WORLD	3.11	4.48	92.41	7.59
	TO OTHERS	12.68	14.55	8.76	TCI
	NET DIRECTIONAL	0.34	-1.51	1.17	12.00
OMN	OMN	83.44	11.45	5.11	16.56
	GCC	11.09	82.66	6.24	17.34
	WORLD	1.71	4.55	93.74	6.26
	TO OTHERS	12.80	16.00	11.36	TCI
	NET DIRECTIONAL	-3.76	-1.34	5.10	13.39
BAH	BAH	91.56	6.33	2.11	8.44
	GCC	5.61	87.53	6.86	12.47
	WORLD	2.00	4.72	93.28	6.72
	TO OTHERS	7.61	11.05	8.96	TCI
	NET DIRECTIONAL	-0.83	-1.41	2.25	9.21

Tax pressure and governance in a spatially interdependent world: international evidence and Morocco's positioning

Mounir ATLASSI

FSJES-Salé, Mohammed V University, Morocco
atlassmounir@gmail.com

Mohamed KARIM

FSJES-Salé, Mohammed V University, Morocco
xmedkarimx@gmail.com

Ilham DKHISSI

International University of Rabat, Morocco
ilham.dkhissi@uir.ac.ma

Abstract. *This paper examines relationships between governance and tax pressure across 118 countries using country-level averages for 2000–2023. Combining Moran's I, an economic-proximity matrix, and cross-sectional regressions, it identifies significant positive clustering in both variables, stronger for taxation. Governance is positively and significantly associated with tax pressure in baseline and spatially augmented specifications. The spatial term absorbs residual dependence, while Morocco records tax pressure above its baseline-predicted level without being an outlier. The findings highlight institutional quality and economic interdependence as correlates of fiscal capacity, while avoiding causal interpretations, and support equitable tax-base development in emerging economies.*

Keywords: Tax pressure, Governance, Fiscal capacity, Spatial econometrics, Tax competition.

JEL Classification: H21, H11, C31, H26, H87.

1. Introduction

Over recent decades, globalization has profoundly transformed the environment in which fiscal policies are designed and implemented. The increasing mobility of capital, the expansion of international trade, and the growing importance of multinational enterprises have intensified economic interdependence across countries. Governments consequently operate within a more competitive fiscal environment in which national tax-policy choices may be associated with developments occurring beyond their borders. In this context, taxation is no longer solely a domestic instrument for financing public expenditure but also constitutes an important component of international economic competitiveness and strategic policy interaction.

The literature on tax competition has extensively examined how the mobility of tax bases constrains governments' capacity to raise revenue and may generate downward pressure on the taxation of mobile factors, particularly capital and corporate profits (Tiebout, 1956; Zodrow and Mieszkowski, 1986; Wilson, 1999). At the same time, empirical evidence reveals substantial differences in fiscal performance among countries facing similar economic conditions. These disparities suggest that economic fundamentals alone may not fully explain cross-country differences in fiscal capacity and revenue mobilization.

In this regard, governance quality has emerged as an important correlate of fiscal performance. Strong institutions can enhance the effectiveness of tax administration, facilitate compliance, reduce opportunities for tax evasion and avoidance, and strengthen the credibility of public policies. Countries characterized by higher government effectiveness, stronger regulatory quality, and greater adherence to the rule of law may therefore be better equipped to transform their economic potential into effective public revenue. The relationship between governance and taxation has consequently become a central issue in public finance and development economics.

Despite the expanding literature on governance and fiscal capacity, an important limitation remains. Many empirical studies treat countries as independent units and analyze taxation primarily within a national framework. In an increasingly integrated world economy, however, fiscal and institutional outcomes may also be related to developments occurring in economically comparable countries. Tax competition, policy diffusion, international benchmarking, and institutional convergence provide potential channels through which national outcomes may become interconnected. Ignoring these relationships may therefore result in an incomplete understanding of the international structure of fiscal capacity.

This perspective raises the possibility of spatial dependence in governance and taxation. Countries sharing similar economic characteristics may display comparable institutional arrangements and fiscal outcomes, generating identifiable clusters of governance quality and tax pressure. Although spatial methods have increasingly been applied to regional growth, innovation, and public expenditure, comparatively limited attention has been devoted to the international spatial structure of fiscal capacity and its relationship with governance. Moreover, existing studies often define spatial relationships through geographic proximity, even though international fiscal interactions may also occur among structurally similar economies that are geographically distant.

Against this background, this paper addresses three related questions. First, do governance and tax pressure exhibit significant spatial clustering among economically comparable countries? Second, does governance remain significantly associated with tax pressure after controlling for economic development, trade openness, and productive structure? Third, does this institutional association persist when the tax-pressure pattern observed among economically similar countries is incorporated into the econometric specification? Particular attention is devoted to Morocco, whose fiscal system has undergone important reforms while continuing to face challenges relating to effective tax-base coverage, informality, compliance, and institutional effectiveness.

The empirical analysis covers 118 countries and uses country-level averages calculated over 2000–2023. This aggregation is intended to capture long-term structural characteristics while reducing the influence of short-term fluctuations. Spatial dependence is assessed using Global Moran's I, while the weighting matrix is constructed from economic proximity in terms of income, trade openness, and productive structure. The relationship between governance and tax pressure is subsequently estimated using a cross-sectional OLS model with HC1 heteroskedasticity-consistent standard errors. An exploratory spatially augmented specification incorporating the economic-proximity-weighted average of tax pressure is also estimated to assess the robustness of the governance result.

The findings reveal positive and statistically significant spatial autocorrelation in both governance and tax pressure. Spatial association is stronger for tax pressure than for governance, indicating a more pronounced clustering of fiscal outcomes within the economic-proximity structure. The baseline regressions also show that governance is positively and significantly associated with tax pressure. Importantly, this relationship remains significant at the 1% level after introducing the economic-proximity-weighted average of tax pressure. The augmented specification eliminates the significant spatial autocorrelation detected in the baseline residuals, although it is interpreted as exploratory robustness evidence because the spatial term may be endogenous.

Morocco occupies an intermediate position within the international fiscal and institutional space. Its observed tax pressure exceeds the value predicted by the baseline model, while its standardized residual indicates that the country is not an extreme observation within the sample. This positioning suggests a meaningful capacity for revenue mobilization relative to its observed economic and institutional characteristics, while also highlighting continuing challenges concerning tax-base breadth, informality, compliance, and the distribution of the tax burden.

This paper contributes to the literature in three respects. First, it provides international evidence on the spatial clustering of governance and tax pressure. Second, it constructs an economic-proximity weighting matrix that complements conventional geographic approaches by capturing relationships among structurally comparable economies. Third, it shows that the positive association between governance and tax pressure remains stable after accounting for economic-proximity-based fiscal patterns, while providing a comparative assessment of Morocco's position within this international structure.

The analysis does not claim to establish causal effects or directly demonstrate the mechanisms of tax competition. Rather, it identifies systematic cross-country associations that are consistent with international fiscal interdependence and provides a basis for further investigation using structural spatial models and annual panel data.

The remainder of the paper is organized as follows. Section 2 reviews the theoretical and empirical literature on tax competition, governance, and fiscal capacity. Section 3 presents the data, methodology, and empirical strategy. Section 4 reports the empirical results and Morocco's positioning. Section 5 discusses the findings and their policy implications. Section 6 concludes.

2. Literature Review

Economic globalization has profoundly transformed the environment in which fiscal policies are designed and implemented. The increasing mobility of capital, foreign direct investment, and multinational profits has intensified fiscal competition among countries, reducing governments' autonomy in tax-policy design. In this context, tax competition has become a central issue in public finance, as jurisdictions seek to attract mobile tax bases while preserving their capacity to generate public revenues.

The theoretical foundations of tax competition originate from the seminal contribution of Tiebout (1956), who emphasized the role of factor mobility in shaping fiscal outcomes. This framework was subsequently formalized by Zodrow and Mieszkowski (1986) and Wilson (1999), who modeled tax competition as a strategic interaction among governments. Their analyses suggest that competition for mobile capital may generate downward pressure on tax rates and lead to suboptimal fiscal equilibria. Empirical studies have largely confirmed the existence of tax-competition effects, particularly in relation to corporate taxation and investment decisions (Devereux et al., 2008; De Mooij and Ederveen, 2008).

More recently, the literature on Base Erosion and Profit Shifting (BEPS) has highlighted the growing ability of multinational enterprises to exploit differences across tax systems, thereby increasing the challenges faced by national tax authorities (OECD, 2015; Beer et al., 2020). Empirical research identifies transfer pricing, intra-group profit shifting, and the location of intangible assets as important channels through which multinational enterprises respond to international tax differentials (Bartelsman and Beetsma, 2003; Huizinga and Laeven, 2008; Dischinger and Riedel, 2011; Dharmapala, 2014; Heckemeyer and Overesch, 2017). These developments reinforce the view that fiscal systems operate within an increasingly interconnected international environment.

Despite these external constraints, countries exhibit substantial differences in revenue mobilization. Economies characterized by similar levels of development often display markedly different tax performance, suggesting that economic fundamentals alone cannot explain variations in fiscal capacity. In this regard, governance quality has emerged as a critical determinant of revenue mobilization. Strong institutions improve administrative efficiency, enhance tax compliance, strengthen enforcement mechanisms, and increase

policy credibility (Besley and Persson, 2013; Moore, 2004). Consequently, countries characterized by higher levels of government effectiveness, regulatory quality, and adherence to the rule of law generally achieve greater fiscal capacity and more stable tax systems.

The concept of fiscal capacity has become a central component of the development and public-finance literature. According to Besley and Persson (2013), fiscal capacity reflects the state's ability to mobilize revenue through effective institutions and administrative structures. Countries with stronger fiscal capacity are generally better positioned to finance public goods, support economic development, and respond to external shocks. By facilitating the provision of productive public expenditure, fiscal capacity may also contribute to long-term economic performance (Barro, 1990). Differences in fiscal outcomes may therefore reflect institutional disparities rather than purely economic constraints.

The literature on fiscal effort further reinforces this argument. Traditionally defined as the ratio of actual tax collection to potential tax capacity, fiscal effort seeks to explain why countries with comparable economic structures generate different levels of tax revenue (Lotz and Morss, 1967). Subsequent studies have shown that institutional quality, administrative performance, and the extent of informality play a central role in shaping fiscal outcomes (Bird et al., 2008; Fenochetto and Pessino, 2013). The persistence of informality constrains effective tax-base coverage and revenue mobilization, particularly in developing economies (La Porta and Shleifer, 2014). Fiscal performance therefore reflects not only economic conditions but also governments' capacity to transform economic potential into effective public revenue.

At the same time, governance itself cannot be viewed exclusively as a domestic phenomenon. Institutional reforms increasingly spread through processes of policy diffusion, international benchmarking, and regulatory convergence. Countries often adopt practices observed in economically comparable economies, generating similarities in institutional arrangements and fiscal outcomes. This suggests that governance and taxation may be associated with broader international interaction mechanisms rather than solely with domestic determinants.

These considerations point to the relevance of a spatial perspective. Spatial econometrics provides a framework for analyzing situations in which observations are interconnected and potentially influenced by neighboring or comparable units (Anselin, 1988). Applied to public finance, spatial methods make it possible to identify whether governance quality and tax pressure exhibit clustering patterns across countries. Positive spatial autocorrelation indicates that comparable countries tend to display similar outcomes, whereas negative spatial autocorrelation indicates that comparable countries tend to display dissimilar outcomes.

Recent studies increasingly emphasize the relevance of spatial dependence in public finance, showing that fiscal policies and institutional outcomes may be associated with strategic interactions among jurisdictions. Such relationships may arise through tax competition, policy imitation, institutional convergence, or common responses to comparable economic constraints, generating clustering patterns that conventional econometric approaches may not adequately capture.

Although spatial approaches have been widely applied to regional growth, innovation, and public expenditure, their application to the governance–taxation nexus remains relatively limited. Existing studies generally rely on geographic proximity to define spatial relationships. However, in an increasingly globalized economy, fiscal competition and institutional convergence may also occur among economically similar countries rather than exclusively among geographically adjacent ones. Countries sharing comparable levels of development, openness, and productive specialization may face similar constraints and adopt comparable fiscal strategies regardless of physical distance.

Despite substantial advances in the literature on tax competition, governance, and fiscal capacity, the spatial dimension of their interaction remains insufficiently explored. Existing studies frequently treat countries as independent observations and rarely examine whether governance and tax pressure exhibit structured relationships among economically comparable economies. Furthermore, the predominant reliance on geographic proximity may overlook the role of economic similarity in structuring fiscal and institutional outcomes.

This study addresses these limitations by examining governance and tax pressure across 118 countries using country-level averages over 2000–2023. By employing a spatial framework based on economic proximity rather than geographic distance, it provides evidence on international clustering in governance and fiscal capacity. In addition, it assesses the conditional association between governance and tax pressure and offers a comparative evaluation of Morocco’s position within the international fiscal landscape.

3. Methodology

This study examines the spatial patterns of governance and tax pressure across countries while assessing Morocco’s position within the international fiscal landscape. Given the increasing importance of tax competition and cross-border economic interactions, fiscal capacity cannot be fully understood without considering the interdependencies among economically comparable countries.

The empirical strategy combines exploratory spatial analysis with cross-sectional econometric analysis. First, global Moran’s *I* statistics are used to determine whether governance and tax pressure exhibit clustering patterns among economically similar countries. Second, a cross-sectional regression is estimated to assess the conditional association between governance and tax pressure while controlling for economic development, trade openness, and productive structure. This combined approach makes it possible to examine both the spatial organization of institutional and fiscal outcomes and the structural factors associated with cross-country differences in revenue mobilization.

The analysis is based on a cross-sectional sample of 118 countries, including developed, emerging, and developing economies. Data are obtained from the World Development Indicators (WDI) of the World Bank and the Worldwide Governance Indicators (WGI). For each country, the variables are calculated as averages over the period 2000–2023. This aggregation attenuates short-term cyclical fluctuations and captures the long-run fiscal,

institutional, and economic characteristics of the countries included in the analysis. The resulting dataset therefore comprises 118 observations, one for each country.

Tax pressure (TF) is measured as the average ratio of total tax revenue to gross domestic product over the study period:

$$TF_i = \frac{1}{T_i} \sum_{t=2000}^{2023} \left(\frac{\text{Tax Revenue}_{it}}{GDP_{it}} \times 100 \right) \quad (1)$$

where TF_i denotes the average tax pressure of country i , and T_i represents the number of annual observations available for that country over 2000–2023.

Governance (GOV) is measured using a composite index constructed from three dimensions of the Worldwide Governance Indicators: Government Effectiveness (GE), Regulatory Quality (RQ), and Rule of Law (RL). The composite indicator is first calculated annually and then averaged over the study period:

$$GOV_i = \frac{1}{T_i} \sum_{t=2000}^{2023} \left(\frac{GE_{it} + RQ_{it} + RL_{it}}{3} \right) \quad (2)$$

Where GE denotes Government Effectiveness, RQ denotes Regulatory Quality, and RL denotes the Rule of Law.

This composite indicator captures governments' institutional capacity to formulate, implement, regulate, and enforce public policies. Its construction gives equal weight to the three institutional dimensions and provides a broad measure of governance quality.

Three control variables are included in the analysis. Economic development is measured by the average logarithm of GDP per capita:

$$LNGDPPC_i = \frac{1}{T_i} \sum_{t=2000}^{2023} \ln(GDPPC_{it}) \quad (3)$$

Trade openness (OPEN) is measured as the average ratio of exports plus imports to GDP:

$$OPEN_i = \frac{1}{T_i} \sum_{t=2000}^{2023} \left(\frac{\text{Exports}_{it} + \text{Imports}_{it}}{GDP_{it}} \times 100 \right) \quad (4)$$

Productive structure (IND) is proxied by the average share of industrial value added in GDP:

$$IND_i = \frac{1}{T_i} \sum_{t=2000}^{2023} \left(\frac{\text{Industrial Value Added}_{it}}{GDP_{it}} \times 100 \right) \quad (5)$$

These variables respectively capture countries' level of economic development, degree of international integration, and sectoral composition.

Spatial dependence is assessed using the global Moran's I statistic:

$$I = \frac{n}{S_0} \cdot \frac{\sum_{i=1}^n \sum_{j=1}^n (x_i - \bar{x})(x_j - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad (6)$$

where n is the number of countries, x_i is the value of the variable under consideration for country i , $\bar{x}$ is the corresponding sample mean, w_{ij} is the spatial weight linking countries i and j , and:

$$S_0 = \sum_{i=1}^n \sum_{j=1}^n w_{ij}$$

The statistic is calculated separately for governance and tax pressure. A positive and statistically significant Moran's I indicates that economically similar countries tend to exhibit comparable values of the variable under consideration. Conversely, a negative and statistically significant value indicates that economically similar countries tend to exhibit dissimilar values.

Statistical significance is evaluated using the standardized Z-statistic:

$$Z = \frac{I - E(I)}{\sqrt{\text{var}(I)}} \quad (7)$$

where $E(I)$ and $\text{Var}(I)$ denote the expected value and variance of Moran's I under the null hypothesis of spatial independence.

Unlike conventional approaches based on geographic distance, this study constructs a spatial weighting matrix based on economic proximity. This choice is motivated by the assumption that tax competition, institutional diffusion, and policy convergence are more likely to occur among countries sharing similar economic structures than among countries that are merely geographically adjacent.

For each country, the economic profile is defined as:

$$Z_i = (y_i, \text{OPEN}_i, \text{IND}_i) \quad (8)$$

Governance and tax pressure are deliberately excluded from the construction of the weighting matrix to avoid mechanically generating spatial similarity in the variables whose spatial patterns are being tested.

The three variables included in the economic profile are standardized before calculating economic distances. This standardization prevents differences in measurement scales from disproportionately influencing the proximity matrix. The Euclidean economic distance between countries i and j is defined as:

$$d_{ij} = \sqrt{(Z_{\text{LNGPPC},i} - Z_{\text{LNGPPC},j})^2 + (Z_{\text{OPEN},i} - Z_{\text{OPEN},j})^2 + (Z_{\text{IND},i} - Z_{\text{IND},j})^2} \quad (9)$$

where the lowercase Z terms represent the standardized values of the corresponding variables.

The initial spatial weights are defined as the inverse of economic distance:

$$W_{ij} = \begin{cases} \frac{1}{d_{ij}}, & i \neq j, \\ 0, & i = j \end{cases} \quad (10)$$

The diagonal elements are set to zero because a country cannot be its own neighbor. The weighting matrix is then standardized by row:

$$w_{ij}^* = \frac{w_{ij}}{\sum_{j=1}^n w_{ij}} \quad (11)$$

This specification implies that the intensity of interaction increases as countries become more similar in terms of income level, trade openness, and productive structure.

To assess the conditional association between governance and tax pressure, the following cross-sectional model is estimated using Ordinary Least Squares:

$$TF_i = \alpha + \beta_1 GOV_i + \beta_2 LNGDPPC_i + \beta_3 OPEN_i + \beta_4 IND_i + \varepsilon_i \quad (12)$$

where α is the constant term, β_1 measures the conditional association between governance and tax pressure, β_2 , β_3 , and β_4 are the coefficients associated with the control variables, and ε_i is the error term. Since the analysis is cross-sectional and based on a single long-run average for each country, no country-specific or time-specific effects are included.

Several diagnostic tests are conducted to assess model reliability. Multicollinearity among the explanatory variables is evaluated using centered Variance Inflation Factors (VIFs). Heteroskedasticity is examined using the Breusch–Pagan test, while residual normality is assessed using the Jarque–Bera test. Because the observations are connected through an economic-proximity weighting matrix, Moran's I is also calculated for the OLS residuals to determine whether unexplained spatial dependence remains after controlling for governance and the structural variables.

When heteroskedasticity is detected, heteroskedasticity-consistent standard errors are employed to ensure reliable statistical inference. If significant spatial dependence remains in the regression residuals, the OLS results are interpreted cautiously and a spatial regression specification is considered as an additional robustness analysis.

4. Results

4.1. Descriptive Analysis and Relationships between Variables

Descriptive statistics provide an initial overview of the main characteristics of the sample and help assess the distributional properties of the variables included in the analysis.

Table 1. *Descriptive Statistics and Normality Tests*

Variable	Mean	Std. Dev	Min	Max	Jarque-Bera	Prob.
GOV	0.070	0.945	-1.654	2.036	6.58	0.037
TF	15.705	6.481	0.375	32.902	2.57	0.276
LNGDPPC	8.639	1.465	6.049	11.556	4.47	0.107
OPEN	82.260	46.065	23.197	325.656	721.27	0.000
IND	28.245	10.777	7.786	65.147	40.17	0.000

Note: The statistics are calculated from country-level averages over the period 2000–2023 for 118 countries.

Source: Authors' calculations based on World Bank (WDI) and Worldwide Governance Indicators (WGI) data, using EViews.

The descriptive results reveal substantial cross-country heterogeneity in governance quality and fiscal performance. Governance (GOV), the central explanatory variable of the analysis, has a mean value of 0.070 and a standard deviation of 0.945, with values ranging from −1.654 to 2.036. This wide dispersion highlights marked institutional disparities across the 118 countries, particularly in government effectiveness, regulatory quality, and the rule of law.

Tax pressure (TF) averages 15.705% of GDP and also displays considerable variation, ranging from 0.375% to 32.902%. This dispersion reflects significant differences in countries' capacity to mobilize public revenues. The broad ranges observed for trade openness and industrial value added further illustrate the structural diversity of the economies included in the sample. Taken together, these results suggest that differences in fiscal performance cannot be explained solely by levels of economic development and should also be examined in relation to institutional quality and structural characteristics.

The Jarque–Bera test rejects the null hypothesis of normality for GOV at the 5% significance level ($p = 0.037$), indicating a statistically significant departure from normality in the cross-country distribution of governance quality. Normality is also rejected for OPEN and IND, with trade openness exhibiting the strongest departure. By contrast, the null hypothesis of normality cannot be rejected for TF and LNGDPPC. These distributional patterns provide further evidence of the substantial institutional and economic heterogeneity characterizing the countries included in the analysis.

Table 2. *Correlation Matrix and VIFs*

Variables	GOV	TF	LNGDPPC	OPEN	IND	Centered VIF
GOV	1.000	0.533	0.870	0.365	-0.155	5.292
TF	0.533	1.000	0.380	0.233	-0.363	-
LNGDPPC	0.870	0.380	1.000	0.379	0.083	5.154
OPEN	0.365	0.233	0.379	1.000	0.084	1.188
IND	-0.155	-0.363	0.083	0.084	1.000	1.294

Note: Centered VIFs are reported only for the explanatory variables. The dash indicates that no VIF is calculated for TF, the dependent variable.

Source: Authors' calculations based on WDI and WGI data, using EViews.

The correlation matrix reveals a positive association between governance and tax pressure (0.533), indicating that countries characterized by stronger institutional quality tend to exhibit higher levels of fiscal capacity. This finding is consistent with the fiscal capacity literature, which emphasizes the role of effective institutions in strengthening tax administration, improving compliance, and supporting revenue mobilization.

A strong positive correlation is also observed between governance and economic development (0.870). This relationship is theoretically plausible, since higher-income economies generally benefit from stronger administrative, regulatory, and legal institutions. However, its magnitude warrants closer examination because GOV constitutes the central explanatory variable of the analysis and LNGDPPC captures a related but conceptually distinct economic dimension.

The centered VIF values are 5.292 for GOV and 5.154 for LNGDPPC, indicating moderate collinearity between governance quality and economic development. Nevertheless, these values remain below the conventional threshold of 10 generally associated with severe multicollinearity. By contrast, the VIF values for OPEN and IND are close to one, indicating no meaningful collinearity involving these variables. Overall, the diagnostics do not reveal severe multicollinearity requiring the exclusion of any explanatory variable. Retaining both GOV and LNGDPPC remains theoretically justified, as they capture distinct institutional and economic dimensions of countries' fiscal capacity.

4.2. Spatial Structure of Governance and Tax Pressure

Table 3. *Global Moran's I Statistics*

Variable	Moran's I	Z-statistic	p-value
GOV	0.124	3.005	0.0026
TF	0.196	4.634	< 0.001

Note: The statistics are calculated using country-level averages over 2000–2023 for 118 countries and a row-standardized economic-proximity weighting matrix.

Source: Authors' calculations based on WDI and WGI data, using EViews.

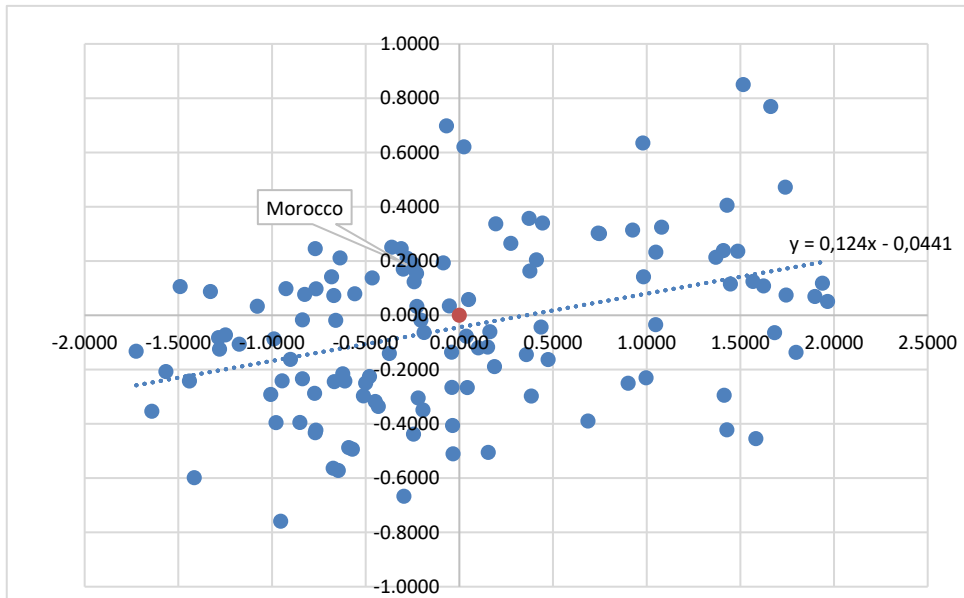
The results reported in Table 3 reveal positive and statistically significant spatial autocorrelation for both governance and tax pressure. Moran's I reaches 0.124 for governance ($Z = 3.005$; $p = 0.0026$) and 0.196 for tax pressure ($Z = 4.634$; $p < 0.001$). The null hypothesis of spatial independence is therefore rejected for both variables. Governance quality and tax pressure are not randomly distributed within the international economic space defined by the economic-proximity weighting matrix. Instead, economically similar countries tend to exhibit comparable institutional and fiscal outcomes.

From an economic perspective, this spatial structuring indicates that institutional quality and fiscal capacity are associated with broader patterns of international interdependence among economically comparable countries. The positive spatial autocorrelation observed for governance is consistent with institutional diffusion, international benchmarking, and convergence among structurally similar economies. However, Moran's I identifies spatial association rather than causality and does not, by itself, establish the mechanisms generating these patterns.

Spatial autocorrelation is stronger for tax pressure than for governance, as indicated by the higher Moran's I value for TF. This result suggests that tax pressure exhibits a more pronounced clustering pattern across economically comparable countries. Such a pattern is consistent with international interaction mechanisms, including tax competition, policy diffusion, and strategic fiscal responses. By contrast, the weaker—although statistically significant—spatial autocorrelation of governance may reflect the more gradual and path-dependent nature of institutional development.

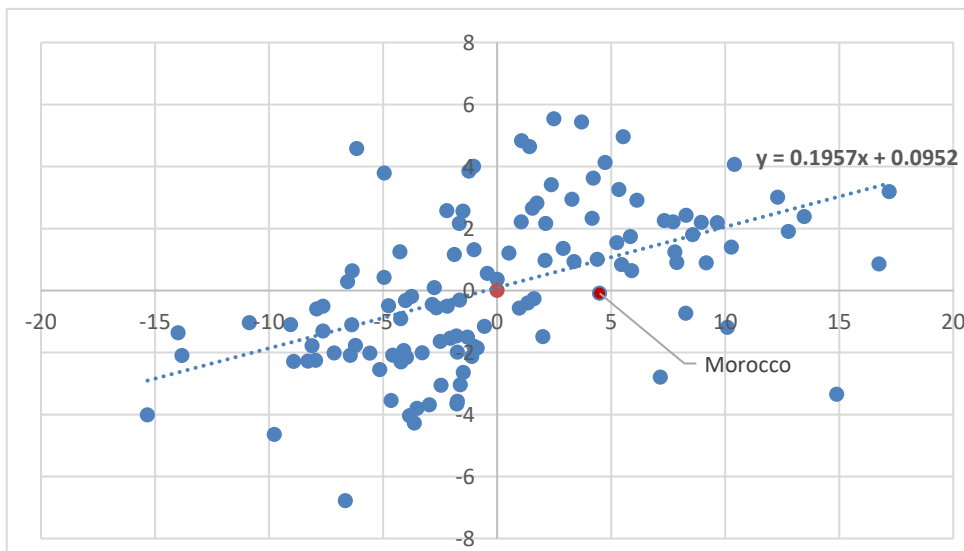
The Moran scatterplots provide graphical confirmation of the positive spatial autocorrelation reported in Table 3. In each figure, the positive slope indicates that countries tend to be associated with economically similar countries displaying comparable values of the variable under consideration.

The four quadrants must be interpreted separately for each variable. In the governance scatterplot, the High–High (HH) quadrant comprises countries with high governance scores associated with economically similar countries that also exhibit high governance scores. The Low–Low (LL) quadrant comprises countries with low governance scores surrounded by comparable economies with similarly low scores. In the tax-pressure scatterplot, the HH and LL quadrants have the same interpretation but refer specifically to tax pressure. In both figures, the High–Low (HL) and Low–High (LH) quadrants identify spatial outliers whose values differ from those of their economically comparable neighbors.

Figure 1. *Moran Scatterplot for Governance*

Note: The horizontal axis represents standardized governance, while the vertical axis represents its spatial lag based on the row-standardized economic-proximity matrix.

Source: Authors' calculations using EViews.

Figure 2. *Moran Scatterplot for Tax Pressure*

Note: The horizontal axis represents standardized tax pressure, while the vertical axis represents its spatial lag based on the row-standardized economic-proximity matrix.

Source: Authors' calculations using EViews.

Overall, the spatial distribution reveals significant clustering patterns in both institutional quality and fiscal capacity. These patterns indicate that governance and tax pressure are embedded in structured relationships among economically comparable countries, thereby supporting the relevance of an economic-proximity approach to understanding fiscal capacity in an increasingly interconnected global environment.

4.3. Cross-Sectional Estimation Results

The relationship between governance and tax pressure is first estimated using a cross-sectional Ordinary Least Squares (OLS) model based on country-level averages over the period 2000–2023. Because the Breusch–Pagan–Godfrey test applied to the baseline specification rejects homoskedasticity at the 5% level ($\text{Obs} \cdot R^2 = 9.988$; $p = 0.0406$), HC1 heteroskedasticity-consistent standard errors are reported.

The presence of statistically significant spatial autocorrelation in the baseline residuals further motivates the estimation of a spatially augmented specification incorporating the economic-proximity-weighted average of tax pressure, denoted WTF. The results of both specifications are presented in Table 4.

Table 4. Cross-Sectional Estimation Results

Variables	Baseline OLS	Spatially Augmented OLS
Constant	23.588*** (5.783)	8.764 (5.822)
GOV	3.839*** (1.126)	3.576*** (1.126)
LNGDPPC	−0.535 (0.751)	−0.466 (0.741)
OPEN	0.014 (0.013)	−0.008 (0.014)
IND	−0.165*** (0.050)	−0.093* (0.050)
WTF	—	0.886*** (0.217)
R ²	0.374	0.453
Adjusted R ²	0.352	0.429
Wald F-statistic	22.022***	24.037***
Prob. (Wald F-statistic)	0.0000	0.0000
Observations	118	118

Note: HC1 heteroskedasticity-consistent standard errors are reported in parentheses. ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels, respectively. WTF denotes the economic-proximity-weighted average of tax pressure. The spatially augmented model is estimated as an exploratory robustness specification rather than as a structural spatial autoregressive model.

Source: Authors' calculations based on World Bank (WDI) and Worldwide Governance Indicators (WGI) data, using EViews.

The baseline results reveal a positive and statistically significant relationship between governance and tax pressure. The coefficient of GOV is 3.839 and is significant at the 1% level ($p = 0.0009$). Conditional on the other explanatory variables, a one-unit increase in the governance index is therefore associated with an increase of approximately 3.84 percentage points in tax pressure. Given the cross-sectional and observational nature of the analysis, this coefficient should be interpreted as a conditional association rather than as a causal effect.

This finding is consistent with the fiscal-capacity literature, according to which stronger institutions facilitate revenue mobilization by improving administrative effectiveness, strengthening regulatory enforcement, promoting taxpayer compliance, and limiting opportunities for tax evasion. Governance therefore emerges as the central institutional correlate of countries' capacity to transform economic activity into public revenue.

Among the control variables, industrial value added has a negative and statistically significant coefficient in the baseline model ($\beta = -0.165$; $p = 0.0013$). This conditional association may reflect cross-country differences in productive structures, sector-specific tax incentives, informality, or the relative contribution of industrial activities to the effective tax base. By contrast, the coefficients of LNGDPPC and OPEN are not statistically significant. Their lack of significance indicates that, after controlling for governance and productive structure, the data do not provide sufficient evidence of an independent association between these variables and tax pressure in this cross-sectional sample. The relatively strong correlation between GOV and LNGDPPC, reflected in their moderate VIF values, may also reduce the precision with which their separate associations are estimated.

The baseline model is jointly significant (robust Wald $F = 22.022$; $p < 0.001$) and explains approximately 37.4% of the cross-country variation in tax pressure. However, Moran's I calculated from its residuals is positive and statistically significant ($I = 0.126$; $Z = 3.042$; $p = 0.0024$). This result indicates that the baseline explanatory variables do not fully capture the dependence generated by the economic-proximity structure.

The spatially augmented specification addresses this issue by introducing WTF, the economic-proximity-weighted average of tax pressure. Its coefficient is positive and highly significant ($\beta = 0.886$; $p < 0.001$), indicating a strong association between a country's tax pressure and that observed among economically similar countries. The inclusion of WTF increases the explanatory power of the model, with R^2 rising from 0.374 to 0.453.

Most importantly, the governance coefficient remains positive and significant at the 1% level in the spatially augmented specification ($\beta = 3.576$; $p = 0.0019$). Its magnitude remains close to that obtained from the baseline model, providing evidence that the central governance result is robust to the introduction of the economic-proximity structure. LNGDPPC and OPEN remain statistically insignificant, while the negative coefficient of IND becomes smaller and is significant only at the 10% level.

Table 5. *Residual Diagnostic Tests*

Diagnostic test	Baseline OLS	Spatially Augmented OLS
Breusch–Pagan–Godfrey Obs* R^2	9.988	9.617
BPG p-value	0.0406	0.0868
Residual Moran's I	0.126	-0.055
Moran Z-statistic	3.042	-1.064
Moran p-value	0.0024	0.2873
Jarque–Bera statistic	11.707	20.275
Jarque–Bera p-value	0.0029	< 0.001

Note: The residual Moran's I statistics are calculated using the row-standardized economic-proximity matrix. The null hypotheses are homoskedasticity for the Breusch–Pagan–Godfrey test, absence of spatial autocorrelation for Moran's I , and residual normality for the Jarque–Bera test.

Source: Authors' calculations using EViews.

The diagnostic results show that the spatially augmented model successfully absorbs the residual dependence detected in the baseline specification. Its residual Moran's I is negative and statistically insignificant ($I = -0.055$; $Z = -1.064$; $p = 0.2873$), meaning that the null hypothesis of spatial independence can no longer be rejected. The Breusch–Pagan–Godfrey test also fails to reject homoskedasticity at the 5% level in the augmented model ($\text{Obs} \cdot R^2 = 9.617$; $p = 0.0868$), although the result remains marginal at the 10% level. HC1 standard errors are therefore retained as a precaution.

The Jarque–Bera test rejects residual normality in both specifications. In the spatially augmented model, this result primarily reflects excess kurtosis rather than pronounced asymmetry. Given the sample size and the use of heteroskedasticity-consistent standard errors, this departure does not invalidate the estimated conditional relationships, but it requires cautious statistical interpretation.

Overall, the findings provide consistent evidence of a positive relationship between governance and tax pressure. This association remains statistically significant after controlling for economic development, trade openness, productive structure, and the tax-pressure pattern observed among economically comparable countries. Nevertheless, because the spatial lag is introduced directly into an OLS framework and may be endogenous, the augmented results should be interpreted as exploratory robustness evidence rather than as definitive estimates of a structural spatial autoregressive process.

4.4. Morocco's Positioning in the International Fiscal Space

The descriptive and spatial evidence places Morocco in an intermediate position within the international fiscal and institutional space. Its profile differs both from economies combining strong governance with high revenue-mobilization capacity and from countries characterized by simultaneously weak institutional quality and low tax pressure. Morocco's position therefore reflects a transitional configuration in which a comparatively developed tax system coexists with continuing institutional and structural constraints.

The baseline cross-sectional model provides an additional perspective on this positioning. Morocco's residual is positive and equal to approximately 4.735 percentage points, indicating that its observed tax pressure exceeds the value predicted by governance, economic development, trade openness, and industrial structure in the baseline specification. The corresponding standardized residual is 0.924, which means that Morocco is not a statistical outlier in the international sample. This positive residual should not be interpreted as a direct measure of tax efficiency or optimal taxation; rather, it indicates that Morocco mobilizes more tax revenue than the baseline model predicts on the basis of the selected structural characteristics.

This result may reflect several features of the Moroccan tax system, including its established revenue-administration framework and its significant reliance on relatively stable domestic tax bases. At the same time, revenue mobilization remains concentrated on formal taxpayers and on comparatively less mobile bases, particularly labor income and domestic consumption. Consequently, a relatively substantial tax pressure does not necessarily imply that the tax burden is evenly distributed or that the full revenue potential of the economy has been achieved.

The econometric results attach particular importance to Morocco's institutional trajectory. Among the structural covariates, governance is the only variable that remains positive and statistically significant at the 1% level in both the baseline and spatially augmented specifications. Across the international sample, higher levels of government effectiveness, regulatory quality, and adherence to the rule of law are therefore associated with stronger revenue mobilization. For Morocco, this finding suggests that strengthening fiscal capacity requires attention not only to statutory tax parameters but also to tax-administration effectiveness, regulatory consistency, taxpayer compliance, and the integration of informal activities into the effective tax base.

The spatial analysis further indicates that fiscal performance should not be interpreted solely through domestic characteristics. Tax pressure exhibits significant clustering among economically comparable countries, while the positive and statistically significant coefficient of WTF shows that, across the sample, national tax pressure is associated with the outcomes observed among structurally similar economies. Accordingly, Morocco's fiscal position should be considered not only in relation to its domestic characteristics but also within the broader economic-proximity structure. These patterns are consistent with policy diffusion, international benchmarking, tax competition, and common responses to comparable economic constraints. However, the analysis identifies statistical associations and does not establish the specific mechanisms or direction of causality underlying these relationships.

Morocco's positive baseline residual and intermediate international position therefore convey a nuanced message. The country displays a meaningful capacity to mobilize tax revenue relative to its observed characteristics, but this performance coexists with challenges concerning the breadth of the tax base, the distribution of the tax burden, informality, compliance, and institutional effectiveness. Strengthening governance and broadening the effective tax base consequently appear more appropriate than relying primarily on higher statutory taxation.

Overall, Morocco's positioning illustrates the central conclusion of the analysis: fiscal capacity is shaped jointly by domestic institutional quality, structural economic characteristics, and the broader environment formed by economically comparable countries. Continued improvements in governance can help Morocco consolidate revenue mobilization while promoting a more balanced, inclusive, and sustainable tax system.

5. Discussion

The findings provide statistically significant evidence of positive spatial autocorrelation in both governance and tax pressure. Global Moran's I is positive for governance ($I = 0.124$; $p = 0.0026$) and tax pressure ($I = 0.196$; $p < 0.001$), indicating that these outcomes are not randomly distributed within the international economic space defined by the weighting matrix. Countries with similar levels of income, trade openness, and productive structure tend to display comparable institutional and fiscal characteristics. These results support an interpretation of fiscal capacity that considers both domestic conditions and the broader environment formed by economically comparable countries.

The results also contribute to the literature on international fiscal interactions by showing that relevant relationships need not be limited to geographic proximity. Economic similarity may provide an alternative channel through which fiscal practices, administrative models, and institutional arrangements become associated across countries. Governments frequently compare their tax systems with those of structurally similar economies, respond to comparable competitive constraints, and draw on common international standards. Nevertheless, because the weighting matrix is specifically constructed from economic characteristics, the spatial findings must be interpreted as conditional on this definition of proximity and should not be generalized automatically to geographic or geopolitical relationships.

Spatial dependence is stronger for tax pressure than for governance, as shown by the respective Moran's I values of 0.196 and 0.124. This difference is consistent with the proposition that fiscal outcomes may adjust more readily to external economic conditions, whereas institutional development is generally gradual and path-dependent. Tax systems can respond relatively quickly to changes in trade integration, tax-base mobility, international standards, and the policies adopted by competing economies. Governance structures, by contrast, are shaped by longer-term administrative, legal, and political processes. However, the Moran statistics identify patterns of association rather than the mechanisms or direction of transmission; the observed difference should therefore not be interpreted as direct evidence of tax competition or institutional diffusion.

The cross-sectional estimation results reinforce the central role of governance. In the baseline OLS model, the coefficient of GOV is positive and significant at the 1% level ($\beta = 3.839$; $p = 0.0009$). This indicates that, conditional on economic development, trade openness, and productive structure, a one-unit increase in the governance index is associated with approximately 3.84 additional percentage points of tax pressure. The result is consistent with fiscal-capacity theory, which emphasizes the importance of government effectiveness, regulatory quality, and the rule of law in strengthening tax administration, facilitating compliance, and limiting revenue losses.

Importantly, the governance result remains stable after accounting for the economic-proximity-weighted average of tax pressure. In the spatially augmented specification, the GOV coefficient remains positive and significant at the 1% level ($\beta = 3.576$; $p = 0.0019$), while the coefficient of WTF is also positive and highly significant ($\beta = 0.886$; $p < 0.001$). The similarity of the governance coefficients across the two specifications indicates that the institutional association is not driven solely by the tax-pressure patterns observed among economically comparable countries. The increase in R^2 from 0.374 to 0.453 further shows that incorporating the economic-proximity structure improves the model's explanatory power.

The residual diagnostics provide additional support for this interpretation. The baseline residuals display positive and significant spatial autocorrelation ($I = 0.126$; $p = 0.0024$), suggesting that the conventional OLS model does not fully capture the structure of international fiscal interdependence. After the introduction of WTF, the residual Moran's I becomes statistically insignificant ($I = -0.055$; $p = 0.2873$). The augmented specification therefore absorbs the residual spatial pattern detected in the baseline model. The Breusch–

Pagan–Godfrey test also fails to reject homoskedasticity at the 5% level in the augmented specification ($p = 0.0868$), although HC1 standard errors are retained because the result remains marginal at the 10% level.

The control-variable results require a more nuanced interpretation than initially expected. Neither LNGDPPC nor OPEN is statistically significant in the baseline or spatially augmented specification. Accordingly, the findings do not establish an independent relationship between economic development or trade openness and tax pressure once governance, productive structure, and economic proximity are taken into account. The strong correlation between GOV and LNGDPPC, together with their moderate VIF values, may reduce the precision with which their separate conditional associations are estimated. Their conceptual distinction nevertheless justifies retaining both variables: LNGDPPC represents economic development, whereas GOV captures institutional capacity.

Industrial value added displays a negative and significant coefficient in the baseline model, but its magnitude and significance decline after WTF is introduced. This instability suggests that the relationship between productive structure and tax pressure may partly reflect broader similarities among countries. It may also arise from differences in sectoral tax incentives, informality, industrial composition, and the extent to which industrial activity contributes to the effective tax base. Consequently, the negative coefficient should not be interpreted as evidence that industrial development necessarily reduces fiscal capacity.

The results have important implications for the distribution of the tax burden. Stronger governance may improve revenue mobilization without requiring higher statutory tax rates by increasing administrative effectiveness, expanding effective coverage, and strengthening compliance. By contrast, where institutional capacity remains limited, governments may rely disproportionately on visible and relatively less mobile tax bases, particularly formal employment and domestic consumption. Such reliance can preserve revenue in the short term but may intensify concerns regarding horizontal equity, informality, and the concentration of taxation on compliant taxpayers.

Morocco illustrates this institutional and fiscal tension. Its positive baseline residual of approximately 4.735 percentage points indicates that observed tax pressure exceeds the level predicted by governance, income, openness, and industrial structure in the baseline model. Its standardized residual of 0.924 nevertheless shows that the country is not an extreme statistical observation. This result suggests a meaningful capacity for revenue mobilization relative to the selected characteristics, but it should not be interpreted as a direct measure of tax efficiency, optimality, or equitable burden sharing.

Morocco's fiscal positioning therefore calls for a balanced interpretation. Progress in tax-administration modernization has supported revenue mobilization, yet challenges remain in relation to informality, effective tax-base coverage, compliance, and the concentration of the burden on formal taxpayers. Given the strong and stable association identified between governance and tax pressure, further improvements in government effectiveness, regulatory consistency, and the rule of law may offer a more sustainable route to strengthening fiscal capacity than relying primarily on increases in statutory taxation.

The results should nevertheless be considered in light of several methodological limitations. First, the analysis is based on country-level averages over 2000–2023 and therefore captures long-term structural differences rather than annual dynamics. Although averaging reduces the influence of short-term fluctuations, it does not identify how fiscal and institutional relationships evolve over time. Second, the cross-sectional design establishes conditional associations but does not demonstrate causality. Reverse causality may also exist because stronger fiscal capacity can provide governments with resources to improve institutional quality.

Third, the spatially augmented model introduces WTF directly into an OLS framework. Because this variable is constructed from the tax pressure of economically comparable countries, it may be endogenous. Its coefficient must therefore be interpreted as exploratory evidence of spatial association rather than as a structural spatial spillover estimate. Moreover, the Jarque–Bera test rejects residual normality, principally because of excess kurtosis. HC1 standard errors and the sample size support asymptotic inference, but the estimates should still be interpreted cautiously. Future research could address these limitations through instrumental-variable methods, maximum-likelihood spatial models, spatial two-stage least squares, alternative weighting matrices, and annual or multi-period panel data.

Overall, the analysis indicates that fiscal capacity is associated jointly with domestic institutional quality and the broader environment formed by economically comparable countries. Governance remains the most stable explanatory variable across the estimated specifications, while the spatial findings show that national tax outcomes are embedded in structured international relationships. For Morocco and other emerging economies, sustainable revenue mobilization therefore depends on institutional strengthening, broader and more equitable tax-base coverage, and an informed response to the fiscal practices of structurally comparable countries.

6. Conclusion

This paper examined the relationship between governance and tax pressure across 118 countries, with particular attention to Morocco's position within the international fiscal landscape. The analysis relies on country-level averages over 2000–2023 to capture long-term institutional and economic characteristics. By combining exploratory spatial analysis with cross-sectional econometric estimation, the study assesses both the international structuring of fiscal outcomes and the conditional association between governance and revenue mobilization.

The results reveal positive and statistically significant spatial autocorrelation in governance and tax pressure. Global Moran's I reaches 0.124 for governance and 0.196 for tax pressure, indicating that economically comparable countries tend to exhibit similar institutional and fiscal outcomes. The stronger spatial association observed for tax pressure suggests that fiscal outcomes display a more pronounced international clustering pattern than governance. However, these findings identify statistical association within the economic-proximity matrix and do not, by themselves, establish the underlying mechanisms or direction of transmission.

The cross-sectional estimates highlight the central role of governance. In the baseline model, a one-unit increase in the governance index is associated with approximately 3.84 additional percentage points of tax pressure, conditional on economic development, trade openness, and productive structure. This association remains positive and statistically significant at the 1% level after introducing the economic-proximity-weighted average of tax pressure, with a coefficient of 3.58. Governance therefore emerges as the most stable explanatory variable across the estimated specifications.

The spatially augmented model also provides evidence of a strong positive association between national tax pressure and the tax pressure observed among economically comparable countries. Its inclusion increases the model's explanatory power and eliminates the significant spatial autocorrelation initially detected in the OLS residuals. Nevertheless, because the spatial lag is introduced directly within an OLS framework and may be endogenous, this result should be interpreted as exploratory robustness evidence rather than as a definitive estimate of a structural spatial spillover.

Morocco occupies an intermediate position within the international institutional and fiscal space. Its observed tax pressure exceeds the value predicted by the baseline model by approximately 4.74 percentage points, while its standardized residual of 0.924 indicates that it is not an extreme observation within the sample. This result points to a meaningful capacity for revenue mobilization relative to the country's observed characteristics, but it should not be interpreted as a direct measure of tax efficiency, fiscal optimality, or equitable burden sharing. Challenges relating to informality, effective tax-base coverage, compliance, and the concentration of taxation on formal taxpayers remain important.

From a policy perspective, the findings suggest that sustainable revenue mobilization depends not only on statutory tax parameters but also on institutional capacity. Improving government effectiveness, regulatory quality, tax administration, and the rule of law may strengthen revenue collection while reducing excessive reliance on higher tax rates or a narrow group of compliant taxpayers. For Morocco, broadening the effective tax base, integrating informal activities, improving compliance, and ensuring greater consistency in tax enforcement appear essential for reconciling fiscal capacity with equity and economic competitiveness.

The study is subject to several limitations. The use of period averages captures structural cross-country differences but does not reflect annual dynamics. The cross-sectional design identifies conditional associations rather than causal effects, and reverse causality between governance and fiscal capacity cannot be excluded. Moreover, the results depend on the selected definition of economic proximity, residual normality is rejected, and the spatially augmented OLS specification does not fully address the potential endogeneity of the spatial lag. Future research could extend the analysis using annual panel data, alternative weighting matrices, spatial instrumental-variable techniques, maximum-likelihood spatial models, and dynamic spatial specifications.

Overall, the findings indicate that fiscal capacity is associated with both domestic institutional quality and the broader environment formed by economically comparable countries. Governance remains the most consistent correlate of tax pressure, while the

spatial evidence demonstrates that fiscal outcomes are embedded in structured international relationships. For Morocco and other emerging economies, strengthening institutions and broadening the effective tax base therefore constitute central elements of a more sustainable, balanced, and resilient fiscal strategy.

References

- Anselin, L. (1988). *Spatial Econometrics: Methods and Models*, Kluwer Academic Publishers, Dordrecht. <<https://doi.org/10.1007/978-94-015-7799-1>>
- Barro, R.J. (1990). Government Spending in a Simple Model of Endogenous Growth, *Journal of Political Economy*, Vol. 98, No. 5, S103–S125. <<https://doi.org/10.1086/261726>>
- Bartelsman, E.J., and Beetsma, R.M. (2003). Why Pay More? Corporate Tax Avoidance through Transfer Pricing in OECD Countries, *Journal of Public Economics*, Vol. 87, Nos. 9–10, 2225–2252. <[https://doi.org/10.1016/S0047-2727\(02\)00018-X](https://doi.org/10.1016/S0047-2727(02)00018-X)>
- Beer, S., de Mooij, R., and Liu, L. (2020). International Corporate Tax Avoidance: A Review of the Channels, Magnitudes, and Blind Spots, *Journal of Economic Surveys*, Vol. 34, No. 3, 660–688. <<https://doi.org/10.1111/joes.12305>>
- Besley, T., and Persson, T. (2013). Taxation and Development, in Auerbach, A.J., Chetty, R., Feldstein, M. and Saez, E. (eds.), *Handbook of Public Economics*, Vol. 5, Elsevier, Amsterdam, 51–110. <<https://doi.org/10.1016/B978-0-444-53759-1.00002-9>>
- Bird, R.M., Martinez-Vazquez, J., and Torgler, B. (2008). Tax Effort in Developing Countries and High-Income Countries, *Economic Analysis and Policy*, Vol. 38, No. 1, 55–71. <[https://doi.org/10.1016/S0313-5926\(08\)50006-3](https://doi.org/10.1016/S0313-5926(08)50006-3)>
- De Mooij, R.A., and Ederveen, S. (2008). Corporate Tax Elasticities: A Reader's Guide to Empirical Findings, *Oxford Review of Economic Policy*, Vol. 24, No. 4, 680–697. <<https://doi.org/10.1093/oxrep/grn033>>
- Devereux, M.P., Lockwood, B., and Redoano, M. (2008). Do Countries Compete over Corporate Tax Rates? *Journal of Public Economics*, Vol. 92, Nos. 5–6, 1210–1235. <<https://doi.org/10.1016/j.jpubeco.2007.09.005>>
- Dharmapala, D. (2014). What Do We Know About Base Erosion and Profit Shifting? A Review of the Empirical Literature, *Fiscal Studies*, Vol. 35, No. 4, 421–448. <<https://doi.org/10.1111/j.1475-5890.2014.12037.x>>
- Dischinger, M., and Riedel, N. (2011). Corporate Taxes and the Location of Intangible Assets within Multinational Firms, *Journal of Public Economics*, Vol. 95, Nos. 7–8, 691–707. <<https://doi.org/10.1016/j.jpubeco.2010.12.002>>
- Fenochietto, R., and Pessino, C. (2013). Understanding Countries' Tax Effort, *IMF Working Paper*, No. 13/244, International Monetary Fund, Washington, DC. <<https://doi.org/10.5089/9781484301272.001>>
- Heckemeyer, J.H., and Overesch, M. (2017). Multinationals' Profit Response to Tax Differentials: Effect Size and Shifting Channels, *Canadian Journal of Economics/Revue canadienne d'économique*, Vol. 50, No. 4, 965–994. <<https://doi.org/10.1111/caje.12283>>

- Huizinga, H., and Laeven, L. (2008). International Profit Shifting within Multinationals: A Multi-country Perspective, *Journal of Public Economics*, Vol. 92, Nos. 5–6, 1164–1182. <<https://doi.org/10.1016/j.jpubeco.2007.11.002>>
- La Porta, R., and Shleifer, A. (2014). Informality and Development, *Journal of Economic Perspectives*, Vol. 28, No. 3, 109–126. <<https://doi.org/10.1257/jep.28.3.109>>
- Lotz, J.R., and Morss, E.R. (1967). Measuring Tax Effort in Developing Countries, *IMF Staff Papers*, Vol. 14, No. 3, 478–499.
- Moore, M. (2004). Revenues, State Formation, and the Quality of Governance in Developing Countries, *International Political Science Review*, Vol. 25, No. 3, 297–319. <<https://doi.org/10.1177/0192512104043018>>
- OECD (2015). Measuring and Monitoring BEPS, Action 11 – 2015 Final Report, *OECD/G20 Base Erosion and Profit Shifting Project*, OECD Publishing, Paris. <<https://doi.org/10.1787/9789264241343-en>>
- Tiebout, C.M. (1956). A Pure Theory of Local Expenditures, *Journal of Political Economy*, Vol. 64, No. 5, 416–424. <<https://doi.org/10.1086/257839>>
- Wilson, J.D. (1999). Theories of Tax Competition, *National Tax Journal*, Vol. 52, No. 2, 269–304. <<https://doi.org/10.1086/NTJ41789394>>
- Zodrow, G.R., and Mieszkowski, P. (1986). Pigou, Tiebout, Property Taxation, and the Underprovision of Local Public Goods, *Journal of Urban Economics*, Vol. 19, No. 3, 356–370. <[https://doi.org/10.1016/0094-1190\(86\)90048-3](https://doi.org/10.1016/0094-1190(86)90048-3)>

Management and performance of territorial police structures: a conceptual framework based on indicators

Gabriel Nicolae UȚĂ

Valahia University of Targoviste, Romania
gabriel.uta@yahoo.com

Abstract. *The management of police structures involves the coordination of activities carried out in an environment characterized by uncertainty, operational pressure, legal responsibility and the use of limited public resources. Under these circumstances, the evaluation of management cannot be reduced to the number of interventions, sanctions or criminal acts recorded. The article proposes a multidimensional conceptual framework for identifying and formalizing indicators that could support the managerial evaluation of territorial police structures. The approach has a theoretical and exploratory character: no real data are processed, no concrete units are evaluated and no institutional rankings are formulated. Dimensions regarding the operational context, effectiveness, responsiveness, human resources, use of material and financial resources, quality of service, legitimacy and managerial capacity are discussed. The methodological part develops indicative formulas for expressing indicators, normalizing values, constituting sub-indices and building a synthetic managerial index. The numerical examples are exclusively hypothetical and have the role of clarifying the mathematical mechanism, not of producing empirical findings. The proposed framework can represent a basis for further research, provided that the indicators are validated, the data quality is ensured and the results are interpreted in relation to territorial and institutional specificities.*

Keywords: public management, police management, institutional performance, performance indicators, evidence-based management, legitimacy, composite index.

JEL Classification: H76, H83, M12, C81.

1. Introduction

The management of police structures is a field of research at the intersection of public management, public administration, security studies and organizational performance analysis. Its complexity results from the institutional specificity of police activity, which involves the exercise of public authority, the protection of citizens' rights and freedoms, the maintenance of public order, the prevention and fight against crime, as well as the management of situations characterized by uncertainty, temporal pressure and high levels of responsibility. In this context, the performance of police structures cannot be assessed solely by the number of activities carried out or by the volume of reported operational results. This must be analyzed in relation to the use of resources, the quality of the decision-making process, the coordination capacity, institutional adaptation, managerial stability and the effects produced on the safety of the community.

The economic, social and technological transformations of the last decades have significantly altered the environment in which law enforcement institutions operate. The diversification of forms of crime, the expansion of cybercrime, the mobility of the population, the intensification of the circulation of information and the increase in citizens' expectations have led to the emergence of additional requirements regarding the organization and management of police structures. At the same time, budgetary constraints, staff shortages, fluctuations in human resources and the need to modernize the technical infrastructure force managers to make informed decisions, set clear priorities and constantly monitor the relationship between the objectives set and the results achieved. Police management can no longer be limited, under these conditions, to the hierarchical coordination of activities, but must include tools for strategic planning, performance management, risk analysis and institutional impact assessment.

The particularities of police management are also determined by the strongly regulated and hierarchical character of the organization. Managerial decisions are adopted within a strict regulatory framework, and the autonomy of territorial structures is influenced by centrally established competences, internal procedures and the distribution of resources. The hierarchy ensures unity of command and coherence of interventions, but can lead, in certain situations, to organizational rigidity, delays in the decision-making process or difficulties in adapting to local particularities. From this perspective, one of the main challenges of police management is to strike a balance between institutional standardization and operational flexibility. Territorial structures must comply with general objectives and procedures, but at the same time they must respond to different local problems depending on the demographic, economic, geographical and criminal profile of each community.

Institutional performance appraisal is an essential component of results-oriented management. However, measuring performance in the field of policing entails numerous methodological difficulties. A low level of crime may be the result of effective preventive activity, but it may also be influenced by social, demographic or economic factors that are

not under the direct control of the police. Similarly, the increase in the number of crimes found may indicate a deterioration in the operational situation, but may also reflect a better ability to identify and document the facts. Therefore, the indicators used must be interpreted in context and analysed in an integrated manner. An evaluation based solely on a single indicator may produce incomplete conclusions, as the operational results must be correlated with the available resources, the volume of requests, the degree of staff load and the characteristics of the environment in which each structure operates.

Within this framework, the article aims to analyze the relationship between organizational management and the performance of territorial police structures, with a focus on how planning, coordination, use of resources and monitoring of results contribute to the achievement of institutional objectives. The research starts from the premise that the performance of a police structure is not determined exclusively by the number of police officers or the level of financial resources available, but also by the managerial capacity to organize these resources, set priorities, distribute responsibilities and adapt interventions to the needs of the community. Consequently, the managerial analysis must go beyond the strictly operative perspective and include dimensions related to efficiency, effectiveness, organizational quality, managerial continuity and adaptability.

The main objective of the research is to identify the managerial elements that can influence the performance of police structures and to highlight ways in which the management process can be analyzed in a coherent manner. In order to achieve this objective, several directions are pursued: presenting the particularities of management in public order institutions, examining the concept of performance in the police field, identifying relevant indicators for evaluating the activity and formulating an indicative mathematical model. At the same time, the study aims to highlight the need to move from a predominantly administrative and descriptive approach to an analytical one, in which managerial decisions can be supported by measurable information and interpretations made in relation to the institutional context.

Methodologically, the research combines the theoretical analysis of the concepts specific to public management and performance management with the mathematical formalization of a set of indicators. The theoretical component allows the clarification of the concepts of efficiency, effectiveness, leadership, planning, managerial control, legitimacy and institutional performance. The mathematical component has an exclusively demonstrative role and shows how certain dimensions could be expressed quantitatively in the case of valid data. The article does not process a database, does not estimate statistical models and does not evaluate a concrete police structure. Consequently, formulas must not be confused with empirical results and cannot substantiate, in the form presented, rankings or causal conclusions.

The importance of research derives from the need to develop evaluation mechanisms that provide a more complete picture of the functioning of police structures. The responsible

use of indicators can help identify imbalances, allocate resources more appropriately, prioritise and increase managerial transparency. Indicators should not be transformed into formal objectives pursued independently of operational reality, but should be used as diagnostic and decision-support tools. Their interpretation must be completed by analyzing the organizational context, qualitative evaluations, and examining external factors that may influence results.

2. Literature and identification of relevant indicators

2.1. Performance Management in Police Institutions

The evaluation of management within police structures must be analyzed in the broader context of the transformations produced in public management. The traditional administrative model has emphasised compliance with rules, the exercise of hierarchical authority, procedural control and the performance of tasks set at a higher level. From that perspective, the proper functioning of the institution was assessed mainly by the legality of its activities, organisational discipline and compliance with procedures. However, the development of performance-oriented public management has led to the extension of the evaluation criteria, by including effectiveness, efficiency, quality of services, use of resources and managerial responsibility. In the case of police structures, this change involves moving from simply reporting the activities carried out to examining the results and effects produced on community safety.

Performance management is the process by which institutional objectives are set, activities are planned, resources are distributed, and results are tracked and evaluated. In a police structure, this process is influenced by peculiarities that are not found to the same extent in other public organizations. Policing is carried out in an environment characterized by uncertainty, risk, time pressure, legal responsibility and the need to take quick decisions. In addition, the results of the activity cannot always be separated from the influence of economic, social, demographic and territorial factors. For this reason, police performance cannot be reduced to a single numerical value or to the evolution of a single phenomenon, such as the number of crimes recorded.

The literature highlights the fact that the police produce several categories of public value. A multidimensional normative framework includes reducing crime and victimization, reducing fear and strengthening personal safety, guaranteeing the security of public spaces, fair use of authority, efficient use of resources, quality of services and institutional legitimacy. This perspective leads to the need for a multidimensional evaluation framework, in which the operational results are analyzed together with the use of resources, the quality of the service and the relationship of the institution with the community. Subsequent studies show that the difficulty lies not only in defining indicators, but also in maintaining a balance between what is easy to measure and what is actually relevant to the organization's mission.

Managerial performance should not be confused with operative performance. Operative results refer to the activities and effects produced by the police structure, while managerial performance refers to how management sets priorities, organizes activity, distributes personnel, uses information, and manages available resources. A structure can achieve certain favorable results due to a relatively stable social context, even if managerial practices are insufficiently developed. Conversely, a structure operating in an environment characterised by high crime, intense mobility and staff shortages may perform apparently poorer, although its management makes efficient use of available resources. The evaluation must therefore differentiate the gross results from the results in relation to the actual operating conditions.

2.2. The multidimensional nature of police performance

The activity of the police includes preventing and combating crime, maintaining public order, intervening at citizens' requests, conducting investigations, ensuring road safety, protecting vulnerable people and cooperating with other institutions. These responsibilities are different in terms of the resources required, their duration and their difficulty. Because of this, aggregating them into a single general indicator can lead to the loss of important information. An intervention at a routine event cannot be considered equivalent to managing a complex situation, even if both are recorded in the institution's records as completed activities.

The literature recommends differentiating between inputs, processes, immediate outcomes, and effects. Inputs include personnel, budget, working time, vehicles, equipment, and information infrastructure. The processes describe the actions carried out by the police, such as patrols, controls, interventions and investigations. Immediate results refer to the direct outputs of these activities, such as solving cases, carrying out interventions or carrying out preventive actions. Effects represent changes in safety, public order, citizens' trust or risks in the community. A coherent management system must follow all these levels, avoiding treating the volume of activity as the equivalent of the impact.

This delimitation is relevant because a high volume of activities does not automatically guarantee better results. The increase in the number of controls or sanctions may indicate an increase in activity, but does not directly demonstrate the reduction of risks. In certain situations, effective prevention of a problem can even reduce the volume of interventions needed. Expert analyses show that traditional measurement systems tend to place excessive emphasis on activities that are easy to count, while risk prevention and solving complex problems are more difficult to reflect through conventional indicators.

From a management perspective, a well-targeted preventive activity can be more valuable than a high number of reactive interventions. However, prevention is difficult to assess because it involves observing events that have not occurred. The absence of a negative event can be the result of police measures, but it can also be determined by social, economic

or behavioral changes. Preventive indicators must therefore be interpreted in relation to the nature of the risk, the target group, the implementation period and the characteristics of the area where the intervention has been taken.

2.3. Limitations of traditional indicators

Among the traditional indicators used to describe police activity are the number of reported facts, crime rate, number of interventions, resolution rate, response time and number of sanctions applied. These indicators have the advantage of availability and relative ease of communication. However, each of them reflects only part of the activity and may be influenced by factors that are not under the direct control of management. Benchmarking analyses of measurement systems have shown that organisations continue to predominantly use traditional indicators, even when strategic objectives are formulated in a broader manner.

The number of reported facts may increase due to the intensification of the criminal phenomenon, but also as a result of a greater willingness of citizens to report incidents. Greater trust in the police, improved reporting systems or information campaigns can lead to an increase in the number of facts known to the institution. Under these circumstances, statistical growth should not automatically be interpreted as a deterioration in performance. Conversely, a low number of complaints does not necessarily confirm the existence of a high level of safety, as some categories of facts may remain unreported.

The resolution rate is relevant for the investigative activity, but is influenced by the complexity of the cases, the availability of evidence, the behavior of victims and witnesses, the duration of expertise and cooperation with other judicial institutions. Comparing structures that handle different types of cases can lead to erroneous conclusions. A structure that predominantly investigates complex facts may have a lower resolution rate than a structure that handles less difficult cases, without this difference demonstrating the existence of inferior management.

Response time is an important indicator for police services, especially in the case of urgent events. However, its value is influenced by distances, road infrastructure, crew distribution, degree of urbanization and number of simultaneous requests. Response time must be correlated with the urgency of the event and the quality of the intervention, as speed does not independently guarantee the adequacy of the response. An overall average can hide both situations resolved very quickly and delays in critical cases.

The number of sanctions, controls or actions may describe the intensity of the activity, but does not provide sufficient information about the effects produced. In addition, turning these metrics into rigid goals can lead to task orientation that generates quick numerical results. Thus, prevention, complex investigations, collaboration with the community and activities whose effects become visible in the longer term can be neglected. Indicators should support managerial judgment, not replace it.

2.4. Evidence-based management in policing

Evidence-based management involves the systematic use of information to set priorities, choose interventions and evaluate results. In the police field, this concept is associated with the evidence-based policing model. This approach is structured around directing interventions to relevant problems, testing solutions, and tracking how they are implemented. Contemporary formulation emphasizes the need to combine the best available evidence with professional judgment and community values, preferences, and circumstances.

The application of this approach does not mean that managerial decisions should be based exclusively on numerical values. Data must be combined with professional experience, analysis of the local context, legal framework and risk assessment. Indicators can signal a problem, but they don't always explain its causes. For example, the increase in response time can be associated with staff shortages, inadequate distribution of crews, expansion of the territory covered or an increase in the number of requests. In order to establish a managerial measure, it is necessary to identify the relevant factor, not just to observe the change in the indicator.

In this article, evidence-based management is used as a basis for proposing categories of indicators that could support the evaluation of the activity. The article does not carry out the processing of data and does not aim at the statistical verification of the relationships between indicators. The aim is to identify information that could be relevant to the management of a police structure and explain their role in the managerial process. The proposed indicators are therefore conceptual in nature. They are not used to evaluate a real structure, to make a ranking or to draw conclusions about the effectiveness of a particular police unit.

2.5. Context and operational pressure indicators

The first category that should be considered is the context indicators. They do not directly measure managerial performance, but are necessary for the correct interpretation of results. Among the relevant elements are the population served, the area of the territory, the population density, the share of the urban population, the mobility of the population, the number of localities and the economic or social particularities of the area.

Operational pressure can be described by the number of requests received, the number of interventions, the number of facts reported, the volume of cases in progress and the number of events requiring the mobilization of additional resources. These indicators show the level of demand addressed to the police structure. They are important because two structures with a similar number of police officers can operate in very different conditions if one of them handles a much higher volume of requests.

Contextual indicators should be used to avoid direct and simplified comparisons. The results of a territorial structure cannot be evaluated without reference to the characteristics

of the community and the level of operational pressure. A longer response time can be explained by the size of the territory and the dispersion of localities, as a high volume of reported facts can be associated with an urban center characterized by intense mobility and diversified economic activity.

2.6. Operational effectiveness indicators

Effectiveness indicators should track the extent to which the structure is meeting its operational objectives. This category may include the rate of resolution of cases, the proportion of interventions properly completed, the reduction of some categories of risks, the evolution of serious events and the achievement of the objectives set by the action plans.

The resolution rate can provide information about investigative capacity, but it must be differentiated according to the type and complexity of the cases. Indicators on preventive activity should also track not only the number of actions carried out, but also their relevance to the problems identified. The mere organisation of activities does not demonstrate the achievement of effects, in the absence of a clear link between the measure and the risk pursued.

Effectiveness must be separated from the volume of work. The high number of actions, controls or interventions shows that resources have been used, but does not automatically confirm that the objective has been achieved. For this reason, quantitative indicators must be accompanied by evaluations of the quality of the intervention, the difficulty of the cases and the concrete outcome of the work.

2.7. Indicators on responsiveness

The responsiveness reflects the structure's readiness to respond to citizens' requests and manage events within an appropriate timeframe. This category can include the response time, the proportion of urgent requests handled in the established intervals, the average duration of solving some works and the number of requests left uncompleted.

The response time must be interpreted differently depending on the nature of the event. Urgent requests should be analyzed separately from those that do not pose an immediate risk. Also, urban and rural environments cannot be assessed by identical standards without taking into account distances and crew availability.

The responsiveness is directly influenced by the managerial organization. The territorial distribution of personnel, shift planning, vehicle availability and the way requests are prioritized can produce important differences. From this perspective, indicators on the time and degree of completion of requests can provide useful information about the organization of the activity, but do not represent a complete measure of the quality of management.

2.8. Human resources indicators

Human resources are one of the main components of institutional capacity. Relevant indicators may include job occupancy, number of vacancies, distribution of staff across

structures, average workload, staff turnover, absenteeism and participation in training activities.

The occupancy rate provides an overview of the availability of staff, but does not indicate whether the existing scheme is adapted to real needs. A structure can have a relatively high occupancy rate and, at the same time, face a shortage of personnel in certain departments. For this reason, the distribution of human resources is as important as the total number of employees.

Staff loading can highlight imbalances between the number of jobs and the number of available staff. However, the activities do not have the same level of complexity, and a simple average can hide important differences. Vocational training indicators may show interest in skills development, but the number of hours of training must be analysed together with the relevance of the programmes and the possibility of applying the knowledge acquired.

2.9. Indicators on the use of material and financial resources

The management evaluation may also include indicators related to budget utilization, equipment availability, fleet status, access to technologies and costs of activities. These indicators are relevant for assessing management's ability to translate available resources into services and results.

The degree of use of the budget should not be confused with efficiency. Spending the funds in full does not automatically demonstrate that resources have been properly distributed. The analysis should follow the link between expenditure, institutional priorities and the results pursued. Also, the costs can vary depending on the type of interventions, the area of the territory and the technical needs of the structures.

Economic indicators should be used with caution in the field of policing. Cost reduction cannot be an independent objective when it affects the safety of staff, the quality of interventions or the protection of citizens. Efficiency should be understood as the rational use of resources to achieve objectives, not as the mechanical reduction of expenses.

Contemporary literature attaches major importance to the way citizens perceive the work of the police. Trust, satisfaction, procedural fairness and respectful treatment are dimensions that can influence the cooperation of the population with the institution. Research on procedural justice shows that legitimacy is associated not only with the outcome of an intervention, but also with the neutrality of the decision, the possibility of people to express their point of view, the respect shown in the interaction and the trust in the intentions of the authority.

This category can take into account the satisfaction of the people who have come into contact with the police, the number and nature of complaints, the proportion of confirmed complaints, the perception of the fairness of the treatment and the level of trust in the institution. This information can be obtained through questionnaires, interviews or the

analysis of complaints. The methodological analyses highlight the need to use multiple measures to assess the public's perspective and the risk of misinterpretation of general questions on trust or satisfaction.

The number of complaints should not be interpreted in isolation. An increase can be determined by real problems, but also by a better knowledge of the referral procedures. Similarly, a small number of complaints does not automatically demonstrate the existence of a high quality of service. The data must be analysed in relation to the number of interactions, the subject of the complaints and the result of their verification.

In order to evaluate the management itself, the stability of the management, the continuity of strategies, the degree of achievement of objectives, the existence of monitoring mechanisms, the frequency of internal evaluations, the ability to correct deviations and the level of cooperation between departments could be analyzed.

Managerial stability can favor the continuity of projects and the maintenance of a strategic direction, but it should not automatically be interpreted as a favorable outcome. Stable leadership can maintain effective practices, but it can also preserve procedures that no longer correspond to current problems. Stability must therefore be analysed together with adaptability and institutional outcomes.

The degree of achievement of the objectives can only be a useful indicator if the objectives are clearly defined, measurable and realistic. Purely quantitative objectives can encourage a formal orientation towards achieving values, without solving real problems. The evaluation must aim at both the achievement of the objective and its relevance to the needs of the community.

The ability to adapt can be seen in the speed with which management identifies problems, redistributes resources, modifies plans and introduces corrective measures. This dimension is difficult to measure exclusively numerically, which is why it should be completed by analyzing the management documents, procedures and decisions adopted.

The literature analyzed supports the use of a balanced system, which combines operational, organizational, economic and social indicators. There is no single indicator by which the performance of a police structure can be fully assessed. There is also no universal framework that can be applied without adaptation to all structures, regardless of size, attributions and territorial context. The application of the Balanced Scorecard in police organizations aimed precisely to extend the evaluation beyond immediate operational results and to connect indicators with strategic objectives.

A balanced system could include the following dimensions: context and operational pressure, effectiveness of activities, responsiveness, human resources situation, use of financial resources, quality of service, legitimacy and managerial capacity. Indicators should be interpreted together, as the improvement of one dimension can be associated with the deterioration of another.

For example, reducing response time may mean focusing crews on urgent interventions and decreasing the resources available for prevention. Increasing the resolution rate can be achieved by prioritizing simpler cases, while complex investigations receive fewer resources. An evaluation based solely on one of these indicators could reward managerial behavior that does not produce the most important effects for the community.

Indicators should be used as diagnostic tools, not as substitutes for managerial analysis. Their role is to flag differences, trends, and possible issues that require further examination. The interpretation must take into account the context, the quality of the data and the limits of each measure. Therefore, the present research proposes a conceptual framework of indicators, without calculating or testing them on the basis of actual data.

3. Conceptual model and mathematical formalization of the proposed indicators

3.1. Methodological status of formalization

The formalization proposed in this section has a conceptual and demonstrative role. It does not automatically transform police management into a completely quantifiable phenomenon and does not assume that the differences between territorial structures can be explained by a single numerical value. The goal is narrower: to define a coherent architecture through which administrative, operational and organizational information could be ordered, compared and interpreted if there were complete, stable and validated data. In this sense, mathematics is used as a language for disciplining concepts, not as a substitute for professional judgment. Each component must have a clear institutional meaning, an identifiable data source, and a rule of interpretation that distinguishes between context, managerial effort, result, and effect.

The unit of analysis is represented by the territorial structure observed in a constant reference period, which can be monthly, quarterly or annual. The choice of period must avoid mixing phenomena with very different rhythms: the response time may vary from one day to the next, the occupation of positions changes more slowly, and the effects on public trust may become visible only at longer intervals. For this reason, the framework must be regarded as a succession of analytical levels and not as a mechanically executed calculation. Raw values are first defined, then contextualised, favourably or unfavourably oriented, brought on a common scale and only then aggregated. Any empirical application should maintain this order in order to avoid turning purely administrative differences into apparent differences in performance.

An important consequence of this positioning is the separation between mathematical validity and managerial validity. An expression can be impeccably calculated and still give a poor picture of the phenomenon if the selected variables do not capture the mission of the institution, if the denominator is not comparable or if the indicator reacts to external factors. In applied research, validation should include at least three levels: checking the definitions

together with practitioners, testing the statistical properties of the series, and comparing the results with qualitative information about the decisions made during the periods analyzed. This triangulation is especially necessary in law enforcement organizations, where results are produced through the interaction between legal rules, resources, organizational behavior, and community characteristics. Formalization is useful precisely because it makes the hypotheses visible, but it cannot decide on its own whether those hypotheses are correct.

In addition, any outcome should be accompanied by a description of its uncertainty. In the case of administrative data, the uncertainty stems not only from statistical error, but also from changes in procedures, differences in recording and conceptual limits of the indicator. Reporting these limits does not weaken the model; on the contrary, it prevents its use beyond what the information can support and preserves the distinction between measurement, interpretation and managerial decision.

3.2. Indicator architecture and working notation

The framework uses a narrow set of variable families so that the model remains intelligible. The context data include the population served, the area of the territory, the density, the degree of urbanization, the mobility and the particularities of the network of localities. Demand for services data includes referrals, interventions, causes or works under management and events that require the simultaneous mobilisation of several resources. Internal capacity is described by planned and employed personnel, personnel with operational attributions, technical availability of vehicles and equipment, financial resources and actual working time. The results are treated separately from these entries and include the resolution of the works, the meeting of deadlines, the response to requests, the achievement of objectives and the quality of interaction with the public.

Notation should not become more important than definitions. Two databases can use the same label for "case solved", but they can apply different rules regarding the time of registration, stock from previous periods or the reopening of works. Before any calculation, each variable must be accompanied by a technical sheet specifying the numerator, denominator, unit of measurement, source, periodicity, reporting officer and rules for dealing with missing values. The same discipline is required for qualitative variables converted into scores. Managerial stability, internal cooperation or the relevance of professional training cannot be credibly introduced into a model just by assigning a grade; they require observable criteria and reproducible assessment procedures.

To avoid overloading the model, each variable should only be introduced if it brings distinct information and can be interpreted managerially. Highly correlated indicators can describe the same reality in different forms and lead to the implicit doubling of the importance of a dimension. For example, the number of requests, the number of interventions and the volume of calls can evolve almost identically in a given structure, in which case aggregating them without verification would overestimate the operational pressure.

Conversely, two seemingly similar indicators can capture different phenomena if one describes the inputs and the other the completion of activities. Therefore, a conceptual mapping stage is recommended before the final selection, in which each indicator is linked to the managerial question to which it answers. This step reduces the risk of building a data-rich system that is poor in decision-making significance.

An initial inventory of the data should also include the level at which each variable can be observed: structure, compartment, locality, type of event or category of cause. Premature aggregation can hide the focus of a problem in a tight area, while excessive granularity can lead to instability and data protection difficulties. The choice of the level of analysis must, therefore, be linked to the managerial decision pursued. For the distribution of resources, territorial and temporal granularity is useful, while quarterly or annual values may be sufficient for strategic evaluation. The coherence between the question, the unit of analysis and the indicator is a basic condition for interpretation.

3.3. Context and operational pressure

Operational pressure must be treated as a contextual dimension, not as a negative outcome attributed to management. A structure that serves a larger population, a larger territory or an environment with intense mobility may use more resources and record higher response times without this automatically indicating poorer management. From this perspective, the absolute number of notified facts, requests or interventions is not sufficient. The information becomes relevant when it is related to the population, area, operative staff and complexity of activities. At the same time, the comparison must avoid the assumption that all requests have the same weight. A simple event and an intervention involving several crews, trips and subsequent activities cannot be treated as identical work units.

For the conceptual model, the pressure is built from four nuclei: the phenomenon known to the institution in relation to the population, the demand for services in relation to the population, the spatial density of the interventions and the workload in relation to the operative staff. The last component can integrate the causes in progress, the outstanding stock and the urgent requests through complexity coefficients. Territorial and mobility factors are used as contextual amplifiers, because the same numerical load can involve different effort in a compact municipality compared to a county with dispersed localities. The following phrase summarizes these relationships in a single block, without assuming that the weights or coefficients are already empirically validated.

Contextualization can also be developed by grouping structures into comparable categories, without turning the grouping into a ranking. Units with high population density, intense mobility and high volume of nocturnal events can be analyzed separately from predominantly rural structures, where distances and dispersion of localities produce other constraints. Such stratification allows for the use of more realistic landmarks and reduces the risk of a structure appearing to be performing or underperforming just because it is compared to a unit in a radically different environment. At the same time, the pressure must

be followed in the dynamics. A moderate but persistent increase in requests may be more managerially important than an isolated peak produced by an exceptional event. Because of this, the analysis of successive periods and the identification of trends must complement any aggregate value calculated for a single interval.

It is also recommended to identify exceptional events that can temporarily alter the profile of a structure: large-scale events, sporting events, severe weather phenomena, tourist periods or special operations. These episodes should not be removed from the data, as they are part of the actual activity, but they should be tagged. Thus, the researcher can separate the structural, persistent pressure from the episodic pressure and observe whether the organization returns to its usual level after the event has ceased.

$$\Pi_{it} = 100 \left[\alpha_1 \frac{F_{it}/P_{it}}{q_F} + \alpha_2 \frac{S_{it}/P_{it}}{q_S} + \alpha_3 \frac{I_{it}/A_i}{q_I} + \alpha_4 \frac{C_{it} + \rho B_{it} + \eta U_{it}}{E_{it}^{op} q_C} \right] (1 + \lambda_1 \tilde{D}_{it} + \lambda_2 \tilde{M}_{it}), \quad \sum_{r=1}^4 \alpha_r = 1, \quad \alpha_r, \rho, \eta, \lambda_1, \lambda_2 \geq 0.$$

In this expression, the weighting coefficients control the contribution of each source of pressure, and the territorial parameters adjust the interpretation without turning the context into a managerial penalty. Benchmarks should be established prior to empirical application, on a comparable and documented basis.

3.4. Operational effectiveness and achievement of objectives

Effectiveness describes the extent to which the work produces the desired results, but must be separated from the mere volume of work. A high resolution rate can have different meanings depending on the complexity of the case portfolio, and exceeding a quantitative target is not always a positive outcome. If the objective is defined by the number of controls, sanctions or actions, the mechanical tracking of the target can stimulate the production of activity that is easy to report, without the guarantee of a relevant effect for the community. Therefore, the evaluation of effectiveness must combine the final result with the difficulty of the activities and the relevance of the objective. In an applied system, the categories of cases should be grouped according to stable criteria, and the weights of complexity should be substantiated by average working time, resources involved or documented professional evaluations.

Another necessary distinction is that between prevention and reaction. The number of preventive actions cannot alone demonstrate that the risk sought has been reduced, but the structure of the activity can indicate whether the management reserves resources for interventions before events occur. The analysis must follow the link between the identified problem, the population or area targeted, the selected measure and the subsequent development. Routine preventive activities, carried out without a risk-based justification, should not be given the same importance as measures targeting documented problems. Consequently, effectiveness is treated as a combination of the completion of activities, the

achievement of the relevant objectives, the complexity of the portfolio and the ability to produce effects without the incentive of artificial reporting.

The evaluation of the objectives must also include the quality of the process by which they were set. An easy-to-achieve goal does not demonstrate performance, just as a very ambitious goal, formulated without regard to resources and context, can artificially produce a low rate of achievement. In practice, objectives can be classified into mandatory outcomes, improvement targets and exploratory outcomes, with each category having a different interpretation. For mandatory objectives, failure to achieve it may signal a problem of capacity or coordination; for improvement targets, the direction and pace of progress matters; For exploratory goals, organizational learning may be relevant even if the numerical target is not fully achieved. This distinction limits the pressure to transform all activities into achieved/unachieved binary values and preserves the link between measurement and the managerial sense of the goal.

Effectiveness can also be tracked by the persistence of the result. An improvement observed in a single period is not sufficient to conclude that a managerial measure has produced a stable effect. It is preferable to examine several periods before and after the intervention, paying attention to external changes that may explain the evolution. In this way, the framework moves closer to an evidence-based assessment and reduces the risk of attributing conjunctural results to a managerial decision.

3.5. Responsiveness and management of the stock of works

Responsiveness is one of the most visible dimensions of the police service, but also one of the most context-sensitive. The average response time is intuitive, but can be distorted by a small number of extreme events. The median, the distribution by priority classes and the proportion of requests within a reference range provide a more stable picture. Urgent events should be analysed separately and reference intervals should be determined according to the nature of the request and territorial conditions. An aggregate indicator that combines critical cases and non-urgent requests can hide exactly the situations on which management needs to intervene. Also, a quick response is not enough if the intervention is incomplete, inadequate or generates repeated replays of the same problem.

Inventory management complements the response time analysis. A structure can respond quickly to new events, but it can accumulate papers, petitions, or backlogs. For management, both the size of the stock and its age, the share of cases with overruns and the concentration of arrears in certain compartments are important. The stock should be related to the entries in the period and the available resources to avoid simplistic interpretations. A higher volume may reflect a temporary increase in demand, a problem with staff distribution, excessively cumbersome procedures or an increase in complexity. The indicator becomes useful only if it is accompanied by a managerial analysis that identifies the source of the accumulation and corrective measures.

Time analysis can be improved by separating the internal stages of the response. The total time between notification and intervention includes, depending on the system, the interval of retrieval and classification, the time until the allocation of the crew, the actual movement and any delays caused by the unavailability of resources. If all these components are aggregated into a single average, management cannot identify the exact source of the problem. A performance management system should allow for the breakdown of the delay and track separately cases where the standard has been exceeded. The same logic applies to the stock of works: the age, the cause of the exceedance, and the number of transfers between departments can be more useful than the total value. Thus, temporal indicators become tools for diagnosing processes, not simple speed measurements.

For high-priority events, the complete distribution of times is more important than the average. Very fast case rates, near-threshold cases, and severe delays may indicate different problems. A management may have an acceptable average but a small number of critical delays that require immediate analysis. For this reason, reporting should also preserve information about the tail of the distribution, not just a central value.

3.6. Human resources and integrated operational capacity

Human resources should not be analyzed exclusively by the degree of occupation of positions. The approved scheme may itself be insufficient for the actual volume of activity, and the internal distribution may produce a deficit in critical compartments even when the overall employment rate is high. Therefore, the evaluation must follow the personnel actually available, the operative staff on shifts, the differentiated load by types of activities, the fluctuation, absenteeism and access to training. Fluctuation requires separating retirements, transfers, promotions, and unplanned departures, as the managerial effects are different. Professional training must be assessed by relevance, completion and applicability, not by the gross number of course hours.

However, operational capacity results from the interaction of staff with other resources and with the results achieved. A structure with full staff, but with unavailable vehicles or with a large portfolio of outstanding works, cannot be adequately described by the occupancy indicator. For this reason, the second mathematical block brings together five components: complexity-weighted resolution, compliance with priority response intervals, effective occupancy, fleet availability and an inversely oriented component on confirmed complaints. The score is not proposed as a final measure, but as an example of how resources and outcomes can be connected without confusing them.

In the evaluation of human resources, the structure of competences must also be taken into account, not just the numerical number of personnel. Two departments with the same number of positions occupied may have different capacities if one has experienced staff, specialised qualifications and continuity, and the other is in a period of integration of new employees. An applied approach could follow the distribution of seniority, critical skills, reliance on a small number of specialists and the existence of succession plans for sensitive

functions. This information should not be used for the individual assessment of police officers, but for the identification of organisational vulnerabilities. From a managerial perspective, capacity means the existence of the right person, with the right competence, in the right place and at the right time; The mere administrative occupation of a position does not guarantee that this condition will be satisfied.

Exchange planning is a distinct link between workforce and performance. The total availability of staff in a month may be satisfactory, but the distribution by hours and days may not be aligned with the peak moments of requests. Comparing the demand curve with the actual presence makes it possible to identify imbalances that do not appear in the annual indicators. This perspective transforms the human resource from an administrative stock into a capacity that must be allocated temporally according to the risk.

$$K_{it} = 100 \left[\beta_1 \frac{\sum_{c=1}^C \kappa_c Sol_{cit}}{\sum_{c=1}^C \kappa_c Reg_{cit}} + \beta_2 \frac{\sum_{j=1}^J \omega_j 1(t_{jit} \leq \tau_j)}{\sum_{j=1}^J \omega_j} + \beta_3 \frac{E_{it}^{occ}}{E_{it}^{plan}} + \beta_4 \frac{V_{it}^{disp}}{V_{it}^{alloc}} + \beta_5 \left(1 - \frac{R_{it}^{conf}}{\max(R_{it}^{ver}, 1)} \right) \right], \quad \sum_{r=1}^5 \beta_r = 1, \quad \beta_r \geq 0.$$

The integrated score is deliberately illustrative. The components can be kept separately in the reporting, and the aggregation should only be used after checking the quality of the data and the managerial sense of the weights. A high value is not a substitute for case analysis and cannot demonstrate managerial causality.

3.7. Material and financial resources

The analysis of material and financial resources must start from the idea that budget execution is not synonymous with efficiency. Full use of the budget may reflect good planning, but it can also indicate spending funds to avoid losing them at the end of the period. Conversely, sub-execution can be associated with savings, delays in procurement, cancellation of procedures or unavailability of necessary products. The budget indicator must be correlated with the financed objectives, the timing of the acquisitions and the results pursued. Unit costs are only useful when there is a credible accounting delineation and when the activities compared are sufficiently similar. The reduction of the average cost cannot become an autonomous target if it affects the safety of the staff, the legality of the intervention or the quality of the service.

Technical availability provides a more direct link to daily operation. For vehicles, it is preferable to measure the days of availability or the hours in which the resource can actually be used, rather than the mere existence in the inventory. For equipment, the need must be established by categories and functions, and the surplus in one category must not compensate for the lack of essential equipment in another. In a managerial analysis, resources should be tracked through a table that combines allocation, availability, utilization, maintenance cost, and effect on activities. Such an approach makes it possible

to differentiate between a lack of resources and the inefficient use of already existing resources.

Another relevant element is the life cycle of resources. The purchase of a vehicle or equipment produces an apparent increase in capacity, but the managerial value depends on availability, maintenance, compatibility with the other systems and subsequent costs. Investments should be analysed in relation to the period during which the resource can actually be used and the activities it supports. In addition, differences between the approved budget and the resources actually accessible may arise from procurement procedures, delivery times or administrative restrictions that are not fully controlled by the local management. To avoid erroneous conclusions, the model must distinguish between allocation, hiring, payment, and actual use. This separation is essential to identify whether a problem stems from a lack of funding, a delay in implementation, or the misuse of an available resource.

For major investments, the assessment may include a pre- and post-implementation phase. It is not enough to notice that a resource has been purchased; It should be checked whether the downtime has been reduced, whether interventions have become faster or whether certain activities can be carried out with less effort. In the absence of such a link, expenditure remains an input indicator. Managerial analysis aims to transform the input into operational capacity and then into the result, while maintaining the separation between these levels so as not to automatically attribute effects to an acquisition.

3.8. Quality of service, integrity and legitimacy

The quality of the service cannot be reduced to the number of complaints. The absolute number is dependent on the volume of interactions with the public, the accessibility of the notification procedures and the willingness of citizens to complain. More informative is the reporting of complaints to interactions and the separation of those confirmed from those not confirmed or still under verification. Even so, the indicator must be interpreted together with the nature of the complaint, the seriousness of the aspect found and the compartment involved. A communication problem and an integrity breach cannot carry the same weight. For management, the usefulness of the indicator derives from the identification of repetitive patterns and the ability to introduce corrective measures, not from simply reducing the number reported.

Satisfaction, procedural fairness and institutional trust require distinct measuring tools. Questionnaires must use multiple items, have stable sampling rules, and allow comparison over time. Procedural justice can be observed through the person's ability to express his or her point of view, the neutrality of the decision, the respect shown and the perception of the authority's intentions. Trust can include the competence, impartiality, integrity, predictability, and availability of the institution. These dimensions should not be aggregated before checking the internal consistency and how respondents understand the questions. In a mature managerial system, quantitative results are complemented by

qualitative analysis of interactions and examination of cases indicating institutional vulnerabilities.

The measurement of public perception must also distinguish between the general population and people who have had recent contact with the police. The former can evaluate the institution based on the public image, indirect experiences and information from the media space, while the latter can appreciate concrete aspects of the interaction. Both perspectives are relevant, but they answer different questions. For management, it is useful to track the difference between overall trust and satisfaction after contact, as well as to identify groups or areas where perceptions deviate significantly. The results should be interpreted with caution when the questionnaire response rate is low or when the sample excludes important categories. A legitimacy score acquires decisional value only if the collection methodology is stable enough to separate a real change from a simple change in the composition of respondents.

Complaints and questionnaires can be completed by analyzing recurrence. If the same types of dissatisfaction occur repeatedly in a compartment or in a certain category of interactions, the managerial signal is stronger than an isolated increase in the total number. Recurrence may indicate a procedural, training, supervision or communication problem. Consequently, the model must allow the transition from the overall score to the concrete operational cause that can be corrected.

3.9. Management organization and quality

Indicators concerning management itself must describe processes over which management has an identifiable influence. Leadership team stability is relevant because frequent changes can disrupt projects, alter priorities, and reduce continuity of accountability. However, stability is not a positive value in itself: stable leadership can maintain both effective practices and persistent dysfunction. Therefore, the duration of mandates must be correlated with the achievement of objectives, adaptation to new problems and implementation of measures resulting from audit or control. More important than mere continuity is the ability to maintain a strategic direction and, simultaneously, to quickly correct decisions that do not produce the expected results.

Managerial adaptability can be analyzed by the sequence between detecting the problem, interpreting the information, deciding to redistribute resources, implementing the measure and verifying the effect. A high-performance managerial structure is not the one that permanently modifies the plans, but the one that can demonstrate that the change was justified by the data and that the result was followed. Compliance with deadlines for managerial measures, implementation of control recommendations and cooperation between departments provide additional clues. These elements are difficult to synthesize exclusively numerically and must be documented through decisions, plans, minutes, evaluation reports and other organizational traces. The role of formalization here is to organize the evidence, not to eliminate institutional analysis.

The quality of management can also be observed by the way information circulates within the organization. Data-driven decisions require channels through which operational information reaches the management level, is analyzed and returns to the structure in the form of clear priorities. Information bottlenecks can cause delays even when resources are sufficient. Therefore, empirical research could track the frequency of operative meetings, the existence of dashboards, the time between the identification of a problem and the issuance of a measure, as well as the mechanisms by which the effect of the decision is subsequently verified. It is not the number of meetings that is relevant in itself, but the link between information, decision and action. This circuit provides a closer perspective to management than purely operational indicators and can explain why structures with similar resources achieve different results in comparable situations.

A complementary element is the quality of delegation. Excessive concentration of decisions at the top of the structure can produce coherence, but it can slow down the reaction to local problems. Delegation without control mechanisms can instead lead to uneven practices. A managerial evaluation can track what types of decisions are decentralized, what information is needed for approval, and how long the decision-making circuit takes. The goal is not to maximize delegation, but to find a balance between speed, responsibility, and control. This balance can explain differences in adaptability even between structures operating under the same regulatory framework.

3.10. Data quality and comparability rules

Systematic data quality control is required before standardisation. It must aim for completeness, consistency of definitions, stability of series over time, duplications, procedural changes and impossible values. A sudden drop in an indicator can reflect a real improvement, but also a change in the way it is recorded. Similarly, the introduction of a new IT system may temporarily increase the number of complaints or reduce administrative delays without operational management having changed to the same extent. Metadata should be kept along with values, and serial breaks should be marked explicitly. Comparability between structures is credible only when the same rules for data production are applied to all analyzed units.

Missing values should not be automatically replaced with zero, because zero is a substantive information different from the absence of information. Imputation may be justified for certain series, but must be reported and tested by sensitivity analyses. Extreme values must be investigated before any statistical limitation. Some may stem from errors, but others may signal a genuine operational situation, such as a major event, severe staff shortages, or a period of exceptional demands. Removing them automatically would precisely reduce the model's ability to identify important pressures. Therefore, data processing must be documented as part of the methodology, not treated as a neutral technical step.

Data traceability is just as important as its accuracy. For each value used in an indicator, it must be possible to identify the initial source, the transformation rule and the time of extraction. If the data is corrected later, the version used in the calculation must be kept for the reproduction of the result. A register of methodological changes should indicate changes in definition, the introduction of new categories and periods when IT systems have been unavailable. In the absence of this documentation, a difference between two years may reflect a technical change and not a managerial one. For research, reproducibility is an essential condition: another researcher should be able to reconstruct the indicator starting from the same data and rules. This requirement reduces the reliance on informal interpretations and turns the measurement system into an auditable tool.

Data control should include automated checks and human checks. Automatic rules can detect impossible values, unusual jumps, or incompatibilities between fields, while professional review can explain whether the deviation is an error or a real event. The separation of the two stages prevents both the mechanical acceptance of wrong data and the discretionary correction of unusual values. Each data intervention must be recorded in order for the final series to remain auditable.

3.11. Normalisation of indicators

The indicators used in the model have incompatible units: minutes, percentages, number of events, costs, days, questionnaire scores, or population-related rates. For aggregation it is necessary to transform them on a common scale, but the choice of method influences the result. Favourable indicators, such as job occupancy or the proportion of interventions carried out within the reference interval, should be targeted so that a higher value corresponds to a better score. Unfavorable indicators, such as response time, outstanding stock or confirmed complaint rate, must be reversed. For indicators with an optimal target, deviation in either direction can be unfavorable. Normalisation should be established before the results are observed in order to reduce the risk of adjusting the rules according to the ranking obtained.

In the case of comparisons between structures, the minimum and maximum limits may come from the observed distribution, from institutional standards or from ex-ante defined thresholds. The use of observed values produces relative scores: one structure can get a high score just because the others perform worse. The use of external standards allows for more stable interpretation over time, but presupposes the existence of justified reference values. Neither option is neutral. Therefore, the methodological report must specify for each indicator the meaning, benchmark and treatment of extreme values. The third mathematical block concentrates in a single expression the three normalization orientations, the formation of subscripts and the geometric aggregation of the six dimensions.

The choice of the reference period for normalisation must also be carefully determined. If the limits are recalculated annually, the scores can remain apparently stable even when all

structures improve or deteriorate simultaneously, because the relative position is preserved. If the limits are fixed for several years, the temporal evolution is easier to follow, but the values can go out of the initial range when the system changes substantially. One solution is to report a relative score and a raw indicator at the same time or related to a stable target. In this way, management can observe both the position vis-à-vis other structures and the absolute progress. Normalization must therefore be seen as a stage of translation of information, not as a transformation that eliminates the need to see the initial values.

For managerial communication, normalized scores must always be accompanied by their meaning. A score of 70 doesn't have the same meaning if it comes from a job occupancy rate, response time, or satisfaction questionnaire. The scoreboard shall allow for a return to the gross value and the baseline. This transparency reduces the risk of scores being treated as abstract units, detached from the institutional phenomenon they represent.

$$N_{kit}^P = \frac{x_{kit} - x_k^{\min}}{x_k^{\max} - x_k^{\min}}, \quad N_{kit}^N = \frac{x_k^{\max} - x_{kit}}{x_k^{\max} - x_k^{\min}}, \quad N_{kit}^T = \max\left(0, 1 - \frac{|x_{kit} - T_k|}{\delta_k}\right).$$

$$D_{dit} = \sum_{k \in d} w_{kd} N_{kit}, \quad \sum_{k \in d} w_{kd} = 1, \quad IM_{it} = 100 \prod_{d=1}^6 \left(\frac{D_{dit}}{100}\right)^{v_d}, \quad \sum_{d=1}^6 v_d = 1.$$

The relationship focuses the entire process: orientation of indicators, transformation on a common scale, aggregation into sub-indices and obtaining a synthetic index. Operational pressure remains outside the final product and is used to contextualize the score, not to automatically diminish it.

3.12. Construction of sub-indices and synthetic managerial index

Aggregation must be carried out in two steps. First, dimensional sub-indices are constructed for operational effectiveness, response capacity, human resources, material and financial resources, quality of service and legitimacy, respectively organization and managerial capacity. This structure prevents numerous indicators in a category from automatically dominating the final result. Operational pressure is kept separate and used for interpretation, not as a direct penalty. Thus, a structure that operates in a difficult environment can be compared with other units without turning the external context into a managerial deficiency. The sub-indices should be published together with the overall index, as the aggregate value does not show where the strengths and vulnerabilities lie.

In the second step, the subscripts can be aggregated arithmetically or geometrically. The arithmetic mean allows for greater compensation between dimensions: a high score on effectiveness can compensate for a poor score on legitimacy. The geometric mean reduces this compensation and penalizes imbalances more sharply. For areas considered non-compensable, such as legality or integrity, a minimum threshold may be introduced, but the threshold must be normatively justified and empirically tested. The index should not be used as an automatic classification or sanctioning mechanism. Its main function is

diagnostic: it signals combinations of results that require investigation, facilitates the tracking of evolution over time and allows the analysis of how managerial changes are reflected in several dimensions simultaneously.

Sub-indices can also be used to build vulnerability profiles. Instead of a single order, structures can be grouped according to the configuration of scores: high pressure and adequate capacity, high pressure and insufficient resources, good effectiveness but low legitimacy, available resources but poor managerial processes. These profiles are more useful for decision-making because they suggest different types of intervention. Some issues require redeployment of staff, others review of procedures, training, investment or interventions on the relationship with the community. The general index can serve as an initial signal, but managerial action must be based on the internal structure of the score. In this sense, the model does not aim to simplify reality down to a single digit, but to organize it in a form that allows the transition from finding to diagnosis and then to corrective measure.

In longitudinal reporting, it is useful to keep the same subscript structure for a sufficient period. Frequent changes in composition can make it impossible to distinguish between changing performance and changing methodology. When an indicator needs to be replaced, the old and new series can be calculated in parallel for a transitional period. This overlap allows the estimation of the methodological effect and maintains the continuity of the managerial interpretation.

3.13. Weighting and governance of the model

Weights are one of the most sensitive choices of a composite index. Equal weights are easy to explain and reduce the impression of artificial precision, but it does not mean that all components have the same institutional importance. An alternative is structured expert consultation, in which the importance of each dimension is independently assessed and then compared. Pair comparisons or other prioritization procedures can also be used, but technical complexity does not automatically guarantee a better decision. Transparency is essential: who set the weights, based on what criteria, with what degree of agreement and what effect they have on the results. If experts express widely dispersed opinions, simple media can hide the lack of consensus.

The governance of the model must also provide for the periodic review of the weights. The mission, risks and availability of data can change, and a set of weights fixed for a period of time can become inappropriate later. However, the revision should not be carried out frequently or opportunistically, as this would compromise temporal comparability. One solution is to establish a review cycle and keep both the series calculated with the original weights and a recalibrated series when the methodology changes. In this way, the system remains adaptable without losing analytical continuity. Decisions on weights, thresholds and definitions should be documented in a methodological protocol separate from the performance reporting itself.

An additional principle is the separation between those who manage the data and those who approve the evaluation rules. If the same structure can change the definitions, weights and values reported, the risk of the system being adjusted to produce a favourable result increases. Methodological governance should provide for distinct responsibilities, documentation of changes and an approval procedure before the new rules enter into force. For research uses, all versions of the methodology must be archived. For internal managerial uses, it may be useful to have a mixed group with operational, legal, statistical and control skills. This diversity reduces the risk that indicators are designed exclusively from the perspective of ease of calculation or, conversely, exclusively from the perspective of operational experience without verification of technical properties.

Before operational use, the model should be tested in a pilot phase. Piloting makes it possible to observe indicators that cannot be fed regularly, definitions interpreted differently and situations in which weights produce unexpected effects. The pilot period should not be used for the formal evaluation of structures, but for the correction of the methodology. After stabilizing the rules, the final version must be frozen for the agreed reporting interval. This discipline limits ad-hoc changes and gives users the opportunity to understand how each result is produced.

3.14. Sensitivity analysis and robustness verification

A composite index is only credible if its results do not change disproportionately to small and plausible variations in methodological assumptions. The sensitivity analysis follows exactly this aspect. Alternative sets of weights, different limits for normalization, alternative treatments of extreme values and different aggregation methods can be tested. If the relative position or conclusion on a structure changes sharply between scenarios, the outcome should be presented as unstable. Conversely, if the values remain close, the index can be used with a higher degree of confidence for descriptive analysis. Sensitivity should not be reduced to a single isolated change; a set of coherent scenarios is preferable, each representing a possible methodological interpretation.

The last mathematical block expresses this logic by recalculating the index for several weighting schemes, measuring the difference from the baseline scenario, and constructing a range of results. Included in the same area is an exclusively demonstrative numerical example with six hypothetical dimensional scores. The values do not belong to any real structure, and the result should not be interpreted as an institutional evaluation. The example is only intended to show that an explicit weighting scheme can be traced back to the aggregate value and that the change in assumptions can be quantified transparently.

Robustness can also be assessed by the controlled exclusion of one component at a time. If the removal of a single indicator radically alters the result, that variable may dominate the model or the other components may provide insufficient information. Another check is to recalculate the score over different periods, to see if the results depend on a particular range. For questionnaire-based indicators, sensitivity may include scenarios on response rate and

sample structure. For administrative data, alternative rules for handling stocks and missing values can be tested. The goal is not to choose the scenario that produces the preferred outcome, but to identify areas where the conclusion is fragile. A robust model is not one that produces exactly the same value under any conditions, but one whose main interpretation remains consistent in the face of reasonable methodological assumptions.

The results of the sensitivity analysis can be communicated through stability ranges and classifications. An index that varies very little between scenarios can be considered robustly descriptive, while a result with high variation must be accompanied by a methodological caveat. It is important that this information is not hidden in technical annexes, as it directly influences how managers can use the score. A seemingly accurate but very weight-sensitive value is less informative than a transparent range that shows the limits of the conclusion. In this way, robustness analysis becomes part of the interpretation and not just a separate statistical check.

$$IM_{it}^{(s)} = 100 \prod_{d=1}^6 \left(\frac{D_{dit}}{100} \right)^{v_d^{(s)}}, \quad \Delta_{it}^{(s)} = IM_{it}^{(s)} - IM_{it}^{(0)}, \quad [IM_{it}^{\min}, IM_{it}^{\max}]$$

$$= [\min_s IM_{it}^{(s)}, \max_s IM_{it}^{(s)}].$$

$$D = (78, 72, 81, 69, 75, 77), \quad v = (0.18, 0.17, 0.16, 0.15, 0.18, 0.16), \quad IM^{(0)} \approx 75.31.$$

The numerical example only shows the calculation mechanism. In real research, the range between the minimum and maximum scenario should be related together with the base value, and any unstable conclusion to reasonable changes in weights should be made with reservations.

3.15. Managerial interpretation of results

The final result should be read along with the dimensional profile and operational pressure. Two structures can achieve the same overall index through very different combinations: one may have high efficacy and low legitimacy, and the other may have moderate but balanced results. For management, these situations require distinct interventions. That's why the recommended reporting is not a one-time ranking, but a table where the overall index is accompanied by sub-indices, trends over time, context indicators, and data quality notes. The comparison with the previous period can be more informative than the comparison between counties, especially when the territorial differences are large and the structures have different operational profiles.

The responsible use of the model also implies avoiding the transformation of indicators into rigid targets. When staff know that the evaluation directly depends on a certain value, the activity can focus on maximizing it, even if the effect on the mission is reduced. The mechanism can be limited by combining volume and quality indicators, auditing definitions, controlled rotation of secondary indicators, and maintaining a professional assessment component. The model is designed as a tool to support decision and research,

not as an algorithm that replaces managerial responsibility. Its empirical application would require validation, piloting, documentation of sources and continuous verification of the behavioral effects produced by the measurement system.

In practice, reporting should be accompanied by attention thresholds and not by absolute labels of the perform/non-performing type. A threshold may signal the need for further analysis when a sub-index declines rapidly, when the difference from its own historical trend becomes large, or when the deterioration of several dimensions occurs simultaneously. Such a rule turns the dashboard into a warning system, not an automatic sanctioning mechanism. Management can then review documents, case distribution, and exceptional events before taking action. This sequence is important for maintaining human responsibility in the decision-making process. Indicators provide consistency and comparability, but the final interpretation must remain connected to the legal framework, the specificities of the mission and qualitative information that cannot legitimately be reduced to a numerical value.

The final interpretation should end with an explicit link between the signal, probable cause, action and the criterion by which the action will be reassessed. Without this loop, the dashboard remains a descriptive tool. Performance management becomes useful when the information leads to a verifiable measure and when the effect of that measure is subsequently observed. Thus, the model can support organizational learning and not just the periodic reporting of values.

4. Discussions

The proposed framework demonstrates that the managerial evaluation of a police structure requires the separation of several levels of analysis. Context and operational pressure describe the conditions under which the structure operates, while effectiveness, responsiveness, resources, and legitimacy describe dimensions over which management can have a direct or indirect influence. Treating all these elements in a single analytical plane could produce incorrect interpretations. A structure with a high volume of requests should not be penalized just because the average response time is longer, just as a structure with a low rate of reported facts should not be automatically considered efficient.

Mathematical formalization brings transparency, as it forces the researcher to define the numerator, the denominator, the orientation of the indicator, the unit of measurement and the aggregation rules. However, mathematical precision does not eliminate conceptual ambiguity. A ratio can be calculated correctly and, at the same time, insufficiently measure the phenomenon being tracked. Therefore, each formalisation block must be accompanied by an institutional justification, a description of the data source and an analysis of the limits.

The synthetic index can facilitate the communication of a general picture, but it involves the loss of details and the introduction of normative choices. Normalisation determines the

relative position of a structure, weights reflect the importance attached to dimensions, and the aggregation method determines the extent to which high performance in one area compensates for poor performance in another. For this reason, the overall index should be published together with the dimensional sub-indices. A single value, presented without structure and without ranges of sensitivity, can create an unjustified impression of objectivity.

5. Conclusions

The management of police structures cannot be evaluated by a single measure or by reporting exclusively on the volume of activity. The specificity of the police mission, the diversity of responsibilities, the influence of the territorial context and the importance of legitimacy require a multidimensional approach. The literature supports the extension of the assessment beyond traditional indicators of crime and the inclusion of resources, quality of service, relationship with the public, integrity and organizational capacity.

The article proposed a framework consisting of indicators of context, effectiveness, response, human resources, material and financial resources, quality of service, legitimacy and management. The four areas of mathematical formalization show how these elements could be expressed mathematically and integrated into subscripts or a synthetic index. Formalization contributes to conceptual clarity, but does not automatically turn indicators into valid performance measures. Validity depends on the quality of the data, the adequacy of the definitions and the interpretation in relation to the context.

The main contribution of the approach is to provide a theoretical architecture that can be used for the design of further empirical research. The article does not process real indicators and does not draw conclusions about the performance of a concrete structure. The numerical examples are exclusively demonstrative, and the synthetic index is presented as a methodological option, not as a validated tool.

Future research could test the availability and comparability of data, validate indicators by consulting experts, and analyze the sensitivity of results to normalization, weighting, and aggregation. Quantitative evaluation should also be complemented by qualitative methods, such as analysis of managerial documents, interviews and examination of decision-making processes. Only the combination of these perspectives can produce a credible picture of how management contributes to the performance and legitimacy of police structures.

The use of indicators in police management can produce behavioral effects. When indicators become rigid targets, staff and management can steer their work towards maximizing the reported metrics. This risk is particularly important in the case of sanctions, controls, interventions and resolution rates. In order to limit formal optimization behaviors, the system must combine volume indicators with quality indicators, avoid mechanical reward of values and maintain the possibility of professional evaluation of the context.

The dimensions of legitimacy, integrity and procedural fairness should not be treated as secondary elements. A structure can achieve seemingly favorable operational results, but it can produce institutional costs through unfair interactions, unequal treatment, or the deterioration of public trust. From this perspective, a balanced managerial model must provide either sufficient weights for these dimensions or minimum thresholds that do not allow the full compensation of poor performance by high results in other areas.

The model remains a conceptual one. For its application, it would be necessary to precisely define the units of analysis, choose the period, validate the sources, harmonize the definitions, treat the missing values and test the robustness. It would also be necessary to consult practitioners to determine whether the selected indicators reflect actual activity and whether the formulas can be fed with available data. In the absence of these steps, the framework should only be used as a tool to structure the scientific discussion.

References

- Hine, K.A., and Davenport-Klunder, K. (2022). From the aspirational to the tangible: Mapping key performance indicators in Australian policing.. *International journal of police science & management.*, 24(2), 382-396.
- Klose, S. (2024). e-defining Evidence-Based Policing.. *Policing: A Journal of Policy and Practice.*, 12(5), 29-38.
- Maslov, A. (2016). *Measuring the Performance of the Police.* 2nd ed. Chicago: Carleton.
- Moore, M.H., and Braga, A.A. (2004). Police performance measurement: A normative framework.. *Criminal justice ethics*, 23(1), 3-19.
- Sherman, L.W. (2013). The rise of evidence-based policing: Targeting, testing, and tracking.. *Crime and justice*, 42(1), 377-451.
- Sparrow, M.K. (2015). Measuring performance in a modern police organization.. *Psychosociological Issues in Human Resource Management.*, 3(2), 17-52.
- Sunshine, J., and. T.T.R. (2003). The role of procedural justice and legitimacy in shaping public support for policing. *Law & society review.*, 37(3), 513-547.
- Tiwana, N., Bass, G., and Farrell, G. (2015). Police performance measurement: An annotated bibliography. *Crime science.*, 4(1), 26-38.
- Tyler, T.R. (2004). Enhancing police legitimacy. *The annals of the American academy of political and social science.*, 539(1), 84-99.
- Wisniewski, M., and Dickson, A. (2001). Measuring performance in Dumfries and Galloway constabulary with the balanced scorecard. *Journal of the Operational Research Society.*, 52(10), 1057-1066.