Economic Structure, Administrative Capacity, and Public Trust in Tax Performance: A Service-Centered Systematic Review across Income Groups, 2015–2025

[a] Vissia Dewi Haptari; [b] Agus Sriyanto; [c] Muhammad Zilal Hamzah; [d] Eleonora Sofilda

[a b] Diploma III Tax Study Program and Diploma III Customs and Excise Study Program, Politeknik Keuangan Negara STAN, South Tangerang, Banten, Indonesia; [c d] Faculty of Economics & Business, Universitas Trisakti, Jakarta, Indonesia

ABSTRAK

Perbedaan rasio pajak terhadap PDB antarnegara mencerminkan variasi struktur ekonomi, kapasitas administrasi, dan kepercayaan publik dalam kerangka kelembagaan yang berbeda. Studi ini menyajikan tinjauan pustaka sistematis periode 2015–2025 untuk menjelaskan determinan rasio pajak lintas kelompok pendapatan negara. Dengan protokol PRISMA, kami mengidentifikasi 150 studi yang memenuhi syarat dan menyusun 90 studi inti untuk sintesis komparatif, dengan pembobotan mutu berdasarkan desain penelitian dan transparansi metodologis. Bukti bermutu tinggi menunjukkan bahwa pendapatan per kapita dan kualitas tata kelola berasosiasi positif dengan rasio pajak, sedangkan inflasi dan rente sumber daya yang dikelola buruk melemahkan upaya pemajakan non-sumber daya; digitalisasi administrasi (e-filing, e-invoicing, analitik berbasis risiko) cenderung meningkatkan kepatuhan, tetapi bergantung pada kapasitas dasar, desain layanan, dan integritas birokrasi. Kajian ini menawarkan kerangka berpusat pada layanan yang mengintegrasikan perspektif kontrak fiskal dan kapasitas institusional dalam siklus "service–trust–revenue" serta "tangga kapasitas pajak", dengan implikasi kebijakan berupa penahapan reformasi: penguatan tata kelola, modernisasi administrasi, dan perluasan basis pajak secara progresif melalui komunikasi transparan dan penegakan digital berbasis risiko.

ABSTRACT

Cross-country variation in tax-to-GDP ratios reflects differences in economic structure, administrative capacity, and public trust under diverse institutional settings. This study provides a PRISMA-guided systematic review (2015–2025) to explain the determinants of tax performance across income groups. We identify 150 eligible studies and compile a core set of 90 for comparative synthesis, applying quality weights based on research design and methodological transparency. Higher-quality evidence links higher income per capita and stronger governance to higher tax ratios, while inflation and poorly managed resource rents weaken non-resource tax effort; digital tax administration (e-filing, e-invoicing, and risk-based analytics) is generally associated with higher compliance, but effects depend on baseline capacity, service design, and bureaucratic integrity. The review advances a service-centered framework integrating fiscal-contract and institutional-capacity perspectives into a "service–trust–revenue" loop and a "tax-capacity ladder", with policy implications on reform sequencing: strengthening governance, modernizing administration, and progressively broadening the tax base through transparent communication and risk-based digital enforcement.

INTRODUCTION

Taxation, measured as total tax revenue over GDP, refers to a first-order indicator of fiscal capacity, state capability, and the breadth of the social contract. Persistent gaps in tax-to-GDP ratios between high- and low-income countries shape macroeconomic stability, development finance, and equity. In general, high-income countries sustain higher ratios because broader

formal tax bases, stronger administrations, and mature compliance ecosystems translate activity into revenue more effectively (Gwaindepi, 2021; Piancastelli & Thirlwall, 2020). Low-income countries, by contrast, face narrower bases owing to extensive informality and sectoral composition, alongside governance frictions that depress collections and weaken perceived reciprocity (Mohammed & Tangl, 2023; Muguchu et al., 2021). These structural differences interact with distributional objectives: progressive systems in advanced economies are associated with lower inequality, while reliance on more regressive instruments in many LICs constrains redistribution (Karakotsios et al., 2020). Institutional quality and the technological backbone of public administration further shape outcomes; recent e-government advances can bolster control of corruption and government effectiveness and, in turn, revenue performance (Mahmoud et al., 2025). At the same time, shadow-economy dynamics and exposure to global shocks can destabilize fiscal systems where state capacity is thin (Mazur et al., 2025). Against this backdrop, a systematic synthesis is needed to clarify what drives tax-ratio differentials and which reformable levers might narrow them.

This review contributes on three fronts. First, it develops an integrated service-centred framework that links economic structure, administrative capacity, and public trust to tax performance across income groups. Second, it synthesizes empirical and comparative evidence from studies using a range of advanced methods (e.g., time-series and panel econometrics, structural equation models, and quasi-experimental evaluations) to clarify when findings on tax capacity and tax effort are robust and when they remain fragile across institutional contexts (Gwaindepi, 2021; Piancastelli & Thirlwall, 2020). Third, it maps areas of scholarly consensus and contention and proposes a typology of reform pathways and sequencing—combining governance improvements, administrative modernization (including digital tax administration), and tax-mix rebalancing—that can inform realistic strategies for closing tax-ratio gaps without undermining equity. Throughout, the review engages with findings on governance quality and compliance behavior (Agbanyo et al., 2024; Mohammed & Tangl, 2023), draws on country-specific determinants that illuminate broader mechanisms (Kebede et al., 2024), and incorporates studies of international tax frictions, such as profit shifting and transfer pricing, that erode domestic bases (Robinson & De Beer, 2021; Sebele-Mpofu et al., 2025).

The review covers the period 2015–2025, spanning the post-1990s reform wave, the 2008 global financial crisis, and COVID-19, thereby capturing both structural and cyclical influences on revenue performance. Geographically, it encompasses cross-country and regional studies, with particular attention to Latin America and the Caribbean and sub-Saharan Africa, where revenue-mobilization trajectories and policy experiments are well documented (Gwaindepi, 2021). Within this scope, the synthesis integrates research on tax bases and instruments (e.g., VAT taxable capacity and effort), administrative modernization (e.g., e-filing, e-invoicing, third-party data), governance and corruption, distributional and behavioural responses, and international tax interactions (Muguchu et al., 2021; Robinson & De Beer, 2021). It also incorporates country-level econometric analyses that identify determinants and, where possible, causal channels in specific contexts, alongside studies linking tax policy and inequality (Karakotsios et al., 2020; Kebede et al., 2024). Finally, research on public perceptions and tax morale reflects the growing recognition that compliance is shaped jointly by service credibility and trust in government (Agbanyo et al., 2024; Mohammed & Tangl, 2023). Studies of aid dependence, conflict, and fiscal fragility are included where they bear on domestic revenue mobilization and the measurement of effort (Diarra et al., 2023).

Key concepts are defined for clarity and comparability. The tax ratio is the share of tax revenue in GDP—a composite signal of a state's ability to convert economic activity into public resources (Piancastelli & Thirlwall, 2020). Tax capacity denotes revenue a country could raise given its

economic structure (e.g., income per capita, sectoral composition, trade openness), while tax effort is the ratio of actual collections to predicted capacity (Gwaindepi, 2021; Piancastelli & Thirlwall, 2020). International organizations operationalize these constructs by benchmarking countries against peers with similar fundamentals. In low-capacity environments, however, large informal sectors, fragmented registries, and administrative frictions complicate both measurement and policy inference (Kebede et al., 2024; Muguchu et al., 2021). Tax morale refers to intrinsic willingness to comply, shaped by perceived fairness and the "service-versus-strain" framing of taxation (Mohammed & Tangl, 2023). Governance quality covers control of corruption and government effectiveness; digital government initiatives can reinforce both and indirectly support revenue performance (Mahmoud et al., 2025). The shadow economy, unreported legal production, remains a salient mediator of volatility and fragility in fiscal systems (Mazur et al., 2025). Where relevant, the World Bank income classification is applied to distinguish HICs from LICs and to treat multinational tax challenges as external constraints on domestic bases (Robinson & De Beer, 2021; Sebele-Mpofu et al., 2025).

Within this conceptual and empirical scope, the review addresses the following research questions: (1) How do structural factors—including income levels, sectoral composition, trade openness, and informality—shape cross-country variation in tax ratios? (2) How do governance quality, administrative capacity, and digital tax administration jointly condition tax performance and compliance behaviour across income groups? (3) What combinations and sequences of reforms in economic structure, administrative systems, and tax mixes emerge from the literature as feasible pathways for narrowing tax-ratio gaps while maintaining or improving equity? These questions structure the subsequent analysis, which first elaborates the service-centred framework and review methods, then synthesizes findings across thematic domains, and finally distils implications for policy design and future research.

Literature Review

The conceptual foundations for explaining persistent cross-country differences in tax-to-GDP ratios rest on two complementary lenses that underpin the service-centered framework used in this review: fiscal-contract theory and institutional-capacity theory. Fiscal-contract theory conceptualizes taxation as a reciprocal exchange in which citizens consent to contribute if government credibly delivers public goods and governs fairly; compliance thus depends on perceived reciprocity, procedural justice, and trust (Nkundabanyanga et al., 2017; Umar et al., 2017). Within this perspective, tax morale, the intrinsic motivation to pay taxes, responds to whether taxation is framed and experienced as a service rather than a strain, and to whether collection is embedded in transparent, respectful interactions between taxpayers and the state (Mohammed & Tangl, 2023; Söderström & Wangel, 2023). Institutional-capacity theory, by contrast, emphasizes the state's administrative, legal, and technological capabilities to identify taxpayers, verify bases, enforce rules, and process information at scale; effective capacity broadens the base and converts economic activity into revenue efficiently, while weak capacity invites evasion and erodes legitimacy (Kebede et al., 2024; Schmoll, 2021). Both lenses are mediated by political-economy factors—distributional conflict, inequality, and vested interests—that shape instrument choice and enforcement intensity, as well as by globalization frictions such as transfer pricing and international tax rules that determine the portability of tax bases (Ranjan, 2023; Robinson & De Beer, 2021; Sebele-Mpofu et al., 2025). Measurement frameworks used by international organizations formalize these ideas through the constructs of tax capacity and tax effort, expected revenue given fundamentals versus realized collections, enabling comparative benchmarking across

income groups and regions (Gwaindepi, 2021; Piancastelli & Thirlwall, 2020). At their intersection lie digital government initiatives, blockchain-style transparency tools, and e-tax platforms, which can simultaneously strengthen capacity, lower transaction costs, and reshape the symbolic meaning of paying taxes (Kabir, 2021; Mohammed & Tangl, 2023; Thi Hai Do et al., 2022). From a public-service-management perspective, recent Indonesian evidence shows that specific service-design features and user capabilities of modern e-tax systems matter: e-registration and taxpayer self-efficacy significantly raise compliance, whereas e-filing alone does not; taken together with compliance, these factors improve public service quality in tax administration (Purnamasari & Tahir, 2025).

To make these lenses operational and reduce abstraction, Table 1 maps the core dimensions of the service-centered framework to typical indicators, example variables, and illustrative studies that inform this review.

Table 1.
Operationalization of the service-centered framework: dimensions, indicative measures, example variables, and illustrative studies

Dimension	Typical indicators and measures	Example variables	Illustrative studies
Economic structure	Level of development; openness; sectoral and trade composition	log GDP per capita, trade-to-GDP ratio, shares of agriculture/industry/services	Piancastelli & Thirlwáll (2020); Gwaindepi (2021); Mawejje & Munyambonera (2016)
Governance quality	Control of corruption; government effectiveness; rule of law	governance indices, corruption perception scores, political stability indices	Mohammed & Tangl (2023); Agbanyo et al. (2024); Diarra et al. (2023)
Administrative capacity	Tax administration resources and systems; registry quality; enforcement tools	staff per taxpayer, ICT investment, registry coverage, audit and risk-scoring use	Schmoll (2020); Schoeman & Feria (2019); Kebede et al. (2024)
Trust and tax morale	Perceived fairness, reciprocity, and service quality	tax morale scores, perceived equity, satisfaction with public services	Nkundabanyanga et al. (2017); Umar et al. (2017); Söderström & Wangel (2022)
Digitalization and	Adoption and use of	e-registration, e-filing, e-invoicing, blockchain-	Kabir (2021); Hai et al.

| technology | e-tax systems; data integration; transparency tools | based records, SPBE alignment | (2022); Purnamasari & Tahir (2025); Marthalina et al. (2025) |

Source: Authors' synthesis based on the included literature (2015–2025).

Note: Indicators and variables are illustrative rather than exhaustive; studies are provided as examples informing the synthesis.

Historically, the literature has evolved from structural accounts that linked revenue performance to levels of development, openness, and sectoral composition toward models that foreground governance quality, legitimacy, and technological modernization. Early cross-country work formalized tax capacity/effort decompositions and placed income levels, trade shares, and price stability center stage, typically using cross-sectional or panel regressions. These studies yielded robust associational patterns but offered limited traction on enforcement frictions, citizen perceptions, and subnational heterogeneity, and rarely addressed endogeneity or simultaneity explicitly (Piancastelli & Thirlwall, 2020). From the mid-2010s onward —and especially during 2015–2025—the field absorbed insights from public finance, behavioural economics, and comparative politics, reframing compliance as co-produced by credible service delivery, fairness perceptions, and anticorruption performance (Agbanyo et al., 2024; Mohammed & Tangl, 2023; Nkundabanyanga et al., 2017). This period also featured sustained attention to the shadow economy and informality as determinants of volatility and fragility in revenue systems, and to the distributional consequences of tax mixes amid rising inequality (Karakotsios et al., 2020; Mazur et al., 2025). Administrative modernization accelerated through e-filing, e-invoicing, and third-party data matching, with evidence accumulating on compliance gains—although the strength and durability of these effects vary by baseline capacity, data architectures, and enforcement cultures (Kebede et al., 2024; Thi Hai Do et al., 2022). Complementing these trends, subnational implementation studies highlight that the success of digital public services hinges on policy content, implementer behaviour, inter-agency work networks, target-group participation, and resources; in South Jakarta, services align with the national SPBE framework but still encounter coordination bottlenecks across agencies (Marthalina et al., 2025).

A number of influential contributions anchor these shifts and illustrate both the strengths and the limitations of prevailing methods. Piancastelli and Thirlwáll provide an integrated treatment of the determinants of tax revenue and tax effort across developed and developing countries, helping to standardize capacity/effort measurement and to clarify the role of macro-structural covariates; however, their models still rely largely on reduced-form associations and face the usual challenges of omitted variables and reverse causality (Piancastelli & Thirlwall, 2020). Gwaindepi's comparative analysis of domestic revenue mobilization across sub-Saharan Africa and Latin America foregrounds regional heterogeneity, institutional quality, and the investment climate as conditioning variables, but also underscores data limitations and difficulties in capturing informality (Gwaindepi, 2021). On the compliance side, Nkundabanyanga et al. articulate how perceived effectiveness and fairness shape taxpayer attitudes in a developing-country setting using survey-based models, while Umar et al.'s interviews explicate "tax boycotts" as a breakdown of the fiscal social contract (Nkundabanyanga et al., 2017; Umar et al., 2017). Mohammed and Tangl's "service-versus-strain" framing operationalizes tax morale and

links equity perceptions to willingness to pay, while still relying on cross-sectional data (Mohammed & Tangl, 2023). Agbanyo et al. provide rare firm-level evidence on how perceived political corruption depresses compliance intentions, reinforcing the centrality of governance but raising questions about unobserved firm characteristics (Agbanyo et al., 2024). From a public-service-management perspective, street-level enforcement capacity and service encounters are pivotal: Schmoll documents the limits of frontline capacity and the corrosive effects of broken public-service expectations among collectors, whereas Söderström and Wangel show how accessible public channels and respectful state–citizen encounters institutionalize compliance (Schmoll, 2021; Söderström & Wangel, 2023). Technology-focused work—from Kabir's analysis of blockchain adoption to Hai et al.'s evidence on e-tax attitudes and uptake—specifies mechanisms through which digitalization can lower discretion and foster efficiency, though experimental or quasi-experimental designs remain the exception rather than the rule (Kabir, 2021; Thi Hai Do et al., 2022). Policy–administration alignment likewise features prominently in VAT fraud mitigation, where instrument design must be married to administrative feasibility (Schoeman & de la Feria, 2019).

At the external frontier, Robinson and Beer revisit corporate income tax determinants under incentive competition, Mpofu et al. detail transfer pricing audit constraints, and Ranjan synthesizes the implications of international tax law for low- and middle-income countries (Ranjan, 2023; Robinson & De Beer, 2021; Sebele-Mpofu et al., 2025). Complementary strands connect fiscal structure to social outcomes, including health financing under alternative tax regimes and the political economy of optimal taxation for poverty reduction, while country-specific causality analyses illuminate how sectoral growth and public expenditure shape revenue (Kanbur et al., 2018; Kebede et al., 2024; Mawejje & Francis Munyambonera, 2016; Reeves et al., 2015). Adding to this canon, Purnamasari and Tahir—using SEM-LISREL on 535 respondents in West Java—show that e-registration and self-efficacy are significant compliance drivers and that, together with compliance, they raise public service quality, while e-filing alone is not a significant predictor (Purnamasari & Tahir, 2025). At the meso-level of service implementation, Marthalina, Dione, and Saribulan identify the policy-content, implementer behaviour, network interactions, participation, and resource bundle—within the SPBE architecture—as decisive for digital service performance, and diagnose coordination gaps (e.g., with UP PTSP) as binding constraints (Marthalina et al., 2025).

Despite convergences, significant debates frame the field and reveal where evidence remains weak. A first controversy concerns identification: while many studies associate digital reforms with higher compliance, the external validity of these effects is contested given heterogeneity in baseline capacity, data architectures, and enforcement cultures; quasi-experimental designs exploiting staggered roll-outs or sharp policy discontinuities remain scarce outside a few middle-income settings (Kebede et al., 2024; Thi Hai Do et al., 2022). Micro-evidence suggests that the returns to digitalization are mediated by implementation determinants and network coordination: significant gains from e-registration and self-efficacy can coexist with weak e-filing effects, and SPBE-aligned platforms may still underperform when subdistrict authorities face limited mandates and fragmented inter-agency linkages (Marthalina et al., 2025; Purnamasari & Tahir, 2025). A second debate centres on sequencing: whether to prioritize administrative modernization and risk-based enforcement or to invest first in repairing the fiscal contract by improving service delivery and transparency. Evidence suggests complementarities, but the optimal bundle likely varies by income level, legitimacy baselines, and the structure of informality

(Agbanyo et al., 2024; Mohammed & Tangl, 2023; Nkundabanyanga et al., 2017; Söderström & Wangel, 2023). Third, scholars dispute the magnitude—and even the measurement—of tax capacity and effort where informality is pervasive, registries are fragmented, and national accounts are noisy; mismeasurement can overstate "low effort" in LICs and misguide policy conditionality (Gwaindepi, 2021; Muguchu et al., 2021; Piancastelli & Thirlwall, 2020). Fourth, distributional controversies revolve around fairness and incidence: studies relate higher tax-to-GDP ratios to lower inequality where progressive instruments are administratively feasible, but LICs often rely on regressive consumption taxes and incentive regimes that catalyse a "race to the bottom," raising questions about the equity–efficiency trade-off (Karakotsios et al., 2020; Ranjan, 2023; Robinson & De Beer, 2021). A fifth debate concerns harmonization and public acceptability—especially for environmental and sin taxes—where cultural norms, trust levels, and crisis conditions shape willingness to pay; legitimacy shocks can durably depress compliance (Elliott et al., 2020; Lois et al., 2019; Muhammad et al., 2022; Public acceptance of carbon tax implementation, 2023). Finally, enforcement controversies highlight constraints of street-level implementation—broken career expectations, weak monitoring, and resource scarcity can neutralize well-designed policies—underscoring the operational core of public service management (Schmoll, 2021). In short, "what works" is conditional on service-design variables (e.g., registration pathways, user efficacy) and on governance of inter-agency interfaces at the front line (Marthalina et al., 2025; Purnamasari & Tahir, 2025).

These strands shape the direction of this review in three concrete ways. First, a fiscal-contract lens structures our interpretation of empirical findings on compliance, tax morale, and perceived fairness, foregrounding studies that connect visible service delivery and anticorruption gains to changes in willingness to pay; in synthesizing evidence, we treat trust as both an outcome of prior performance and a mediator that magnifies or mutes the impact of enforcement (Agbanyo et al., 2024; Mohammed & Tangl, 2023; Nkundabanyanga et al., 2017; Söderström & Wangel, 2023). Second, an institutional-capacity lens organizes the assessment of administrative reforms—e-filing, e-invoicing, data integration, risk scoring, and staff professionalization—and their interaction with legal simplicity; we prioritise evidence that speaks to scalability and to the policy–administration symbiosis central to public service management, recognising that legal instruments are only as effective as the systems that administer them (Kabir, 2021; Kebede et al., 2024; Schmoll, 2021; Schoeman & de la Feria, 2019). Operationally, we therefore pay attention to service-design covariates (e.g., registration pathways, user self-efficacy) and subnational implementation conditions (SPBE alignment, inter-agency coordination, participatory feedback channels) when interpreting reported effect sizes (Marthalina et al., 2025; Purnamasari & Tahir, 2025). Third, a comparative political-economy lens situates domestic reform within international constraints—profit shifting, treaty obligations, and investment competition—and within societal preferences over redistribution and growth. This pushes the review to interrogate claims about capacity and effort using regionally disaggregated studies and to attend to heterogeneity within income groups (Gwaindepi, 2021; Ranjan, 2023; Robinson & De Beer, 2021; Sebele-Mpofu et al., 2025).

Taken together, the literature suggests relatively settled patterns on some fronts and open questions on others. Income levels, governance quality, and macroeconomic stability (including low and predictable inflation) are consistently associated with higher tax ratios, while extensive informality, weak administrative capacity, and volatile resource revenues tend to depress

collections (Mawejje & Francis Munyambonera, 2016; Muguchu et al., 2021; Piancastelli & Thirlwall, 2020; Reeves et al., 2015). By contrast, the measurement of informality and wealth, the design of equitable yet administratively feasible tax mixes, and the sequencing of digital and transparency reforms remain contested, with limited causal evidence and substantial heterogeneity across LICs, MICs, and HICs (Elliott et al., 2020; Kanbur et al., 2018; Muhammad et al., 2022; Samusevych et al., 2024); Public acceptance of carbon tax implementation, 2023). These unresolved issues map directly onto the three research questions set out in the introduction: how structural factors shape tax ratios; how governance, administrative capacity, and digitalization jointly condition performance; and how reform bundles and sequencing can narrow gaps without worsening inequality. The following sections therefore move from this conceptual and empirical groundwork to the review design and quality appraisal, before presenting a thematic synthesis of findings and a critical analysis of gaps and priorities for future research.

RESEARCH METHODS

This article is a structured systematic literature review that synthesises comparative evidence on the determinants of tax-to-GDP ratios across income groups for 2015–2025 using a multi-stage, PRISMA-guided approach. The protocol integrates transparent search, explicit eligibility criteria, dual-coded quality appraisal with a quantitative weighting scheme, standardised data extraction, and a structured synthesis that cross-walks macro-structural determinants with governance, administrative capacity, informality, digitalisation, and external constraints. Consistent with best practice in fiscal systematic and scoping reviews, we privilege replicability, data provenance, and cross-study comparability (Elliott et al., 2020; Ventura-Cots et al., 2019).

Search strategy and information sources

We queried indexed scholarly databases—Scopus, Web of Science Core Collection, EconLit, RePEc/IDEAS, and SSRN—alongside institutional repositories likely to contain high-quality grey literature, notably the IMF eLibrary and the World Bank Documents & Reports. Boolean strings cross-walked the dependent variable family ("tax ratio," "tax to GDP," "tax revenue," "tax effort," "tax capacity," "revenue mobilization") with country context ("high income," "low income," "developing," "advanced," HIC, LIC) and determinant classes ("determinants," "drivers," "predictors," "governance," "informality," "administration," "compliance"). Proximity operators were adapted to each database's syntax to balance precision and recall. We applied backward and forward citation tracking to high-influence items, with sensitivity scans targeting governance, informality, and digital administration using terms such as "control of corruption," "tax morale," "e-tax," and "e-invoicing" (Gwaindepi, 2021; Mohammed & Tangl, 2023; Piancastelli & Thirlwall, 2020; Thi Hai Do et al., 2022). Because cross-country tax performance is shaped by policy architecture, we also canvassed corporate taxation, incentive regimes, and profit-shifting studies to capture external constraints (Robinson & De Beer, 2021; Sebele-Mpofu et al., 2025).The overall screening funnel is summarized in Figure 1.

Eligibility criteria and screening

Inclusion criteria admitted empirical or theoretically grounded comparative research—cross-country panels or multi-country regional analyses—published between 2015 and 2025 that examined determinants of tax ratios, capacity/effort, administration, governance, or distributional structure, and that reported sufficient detail on data sources, samples, time frames, and identification strategies. We excluded single-country case studies lacking a comparative lens, opinion pieces, and non-empirical grey literature without methodological exposition, as well as items outside public finance and development domains. To ensure

evidentiary reliability, priority was given to work using recognised international datasets or well-documented administrative/survey data (Aizenman et al., 2019; Castañeda Rodríguez, 2018).

Screening proceeded through deduplication, title–abstract screening, full-text eligibility assessment, and final inclusion based on relevance and methodological adequacy. Regional balance was monitored to avoid overweighting areas with denser publication (e.g., sub-Saharan Africa, Latin America and the Caribbean), while maintaining variation in institutional baselines pertinent to governance and compliance (Gwaindepi, 2021; Muhammad et al., 2022; Söderström & Wangel, 2023). The final dataset comprised 150 studies in the qualitative synthesis and a core set of 90 studies used for more fine-grained comparative inference.

Figure 1.
PRISMA 2020 flow summary
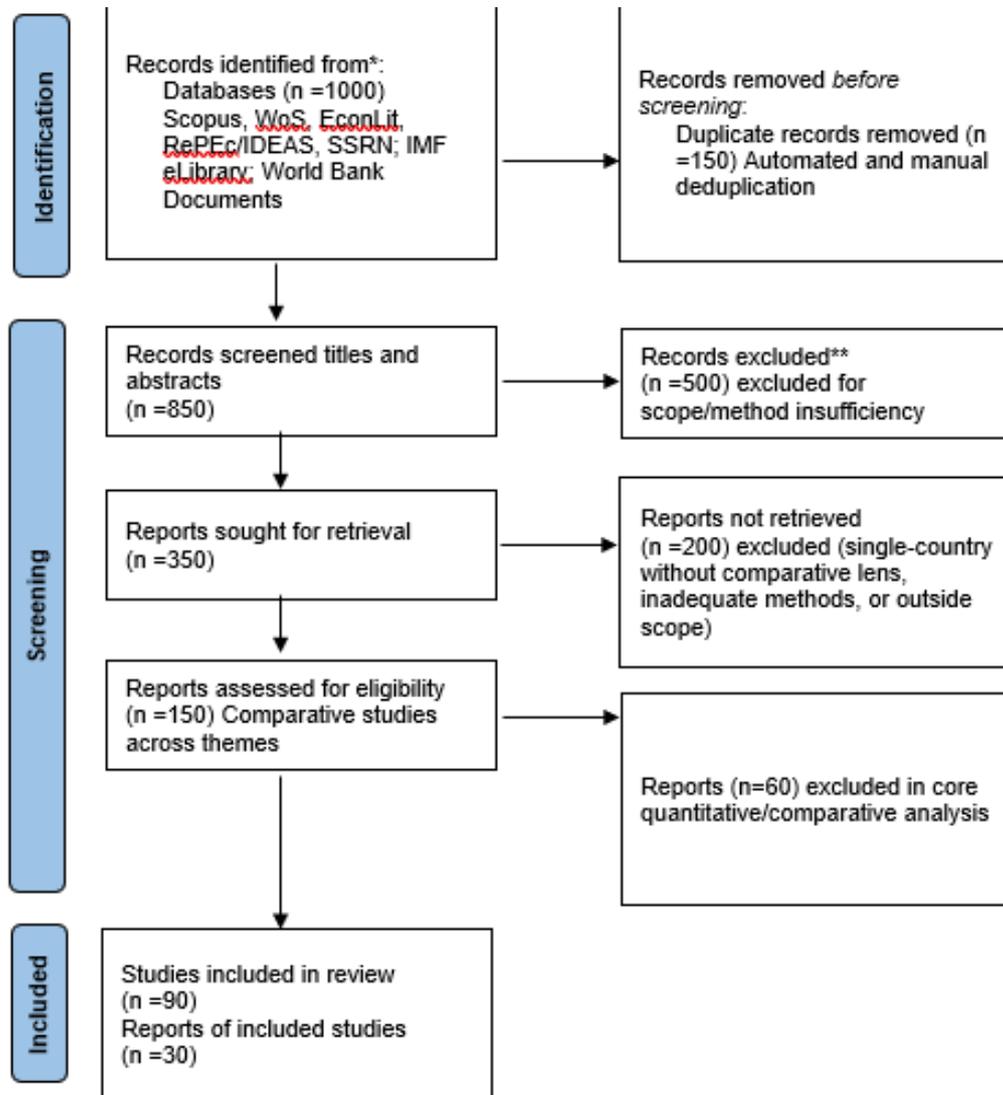
Source: Author's compilation following PRISMA 2020.

Notes. Workflow summarizing records identified, deduplicated, screened (title/abstract), assessed in full text, excluded with reasons, and included in the qualitative (N = 150) and core comparative (n = 90) sets. Reasons for exclusion include scope mismatch, inadequate methods, and insufficient reporting. Counts correspond to the dataset used in this review.

Quality appraisal and narrative weighting

Following recommendations for transparency in taxation reviews (Gnangnon & Brun, 2017; Kebede et al., 2024), two coders independently scored each study on six criteria: (C1) relevance to tax-ratio determinants with an HIC–LIC comparative perspective; (C2) transparency of data and variable definitions; (C3) methodological rigor and identification (e.g., quality of controls, plausibility of causal claims); (C4) cross-study comparability of outcomes and covariates; (C5) replicability (data/code access or method clarity); and (C6) theoretical/empirical contribution. Item anchors were 0 = absent/opaque, 1 = partial/basic, 2 = explicit/robust, producing a 0–12 quality score (Q_i) per study. Disagreements were reconciled by discussion; inter-rater reliability was tracked using weighted Cohen's κ (target ≥ 0.70 before reconciliation).

To structure the synthesis, we derived a simple narrative weight for each study using a two-step scheme. A raw weight $w_i = Q_i/12$ reflects the quality score, adjusted by a design multiplier d_i that captures the relative risk of bias by design class (1.00 = causal/quasi-experimental; 0.90 = well-specified panel/cointegration/SEM; 0.75 = cross-sectional/correlation-heavy or perceptions-only designs). Effective weights were normalized within the sample: $W_i = (w_i \times d_i) / \Sigma_j (w_j \times d_j)$.

These weights are used **solely** to guide the qualitative synthesis: they influence how much interpretive emphasis each study receives in the narrative, in summary tables, and in directional tallies of positive/negative/mixed associations. We do **not** re-estimate pooled regressions or compute meta-analytic effect sizes across studies. Instead, we assign higher narrative weight to studies with stronger identification and clearer reporting, while still incorporating lower-weight evidence where it illuminates mechanisms, implementation conditions, or contexts that are under-represented in the econometric literature. An extract of appraisals for core studies is provided in Table 2.

Table 2.
Quality appraisal scores (extract for core studies)

Study	C1	C2	C3	C4	C5	C6	TQ	Tertile	Design tag	d_i	Raw w	Effective W	Notes
Piancastelli & Thirlwáll (2020)	2	2	2	2	1	2	11	High	Panel (FE)	0.90	0.92	0.082	Transparent data; strong comparability; no experimental variation
Reeves et al. (2015)	2	2	2	2	1	2	11	High	Cross-national modelling	0.90	0.92	0.082	Model-based; external validity depends on assumptions
Aizenman et al. (2019)	2	2	2	2	1	2	11	High	Panel econometrics	0.90	0.92	0.082	Broad coverage; solid identification checks
Robinson & Beer (2020)	2	2	2	2	1	2	11	High	Panel (FE)	0.90	0.92	0.082	Good policy variation; panel controls

Study	C1	C2	C3	C4	C5	C6	TQ	Tertile	Design tag	d_i	Raw w	Effective W	Notes
Mawejje & Munyambonera (2016)	2	2	2	1	1	2	10	Medium	Time series (cointegration)	0.90	0.83	0.074	Country time series; robust but context specific
Kebede et al. (2024)	2	2	2	1	1	2	10	Medium	Time series (cointegration)	0.90	0.83	0.074	Strong diagnostics; single-country limits external validity
Muguchu et al. (2021)	2	2	1	2	1	1	9	Low	VAT capacity–effort	0.90	0.75	0.060	Country-level design; identification limited
Gwaindepi (2021)	2	2	1	2	1	2	10	Medium	Comparative synthesis	0.75	0.83	0.069	Comparative not causal; relies on secondary data
Bérgolo et al. (2023)	2	2	1	2	1	2	10	Medium	Policy synthesis with examples	0.75	0.83	0.069	Narrative synthesis; limited identification
Elliott et al. (2020)	2	2	1	2	1	2	10	Medium	Scoping review	0.75	0.83	0.069	Systematic scope; no causal estimation
Hai et al. (2022)	2	2	1	1	1	2	9	Low	Survey mediation / SEM	0.75	0.75	0.050	Perception-based; risk of common-method bias
Mohammed & Tangl (2023)	2	2	1	1	1	2	9	Low	Mixed-methods survey	0.75	0.75	0.050	Cross-sectional; limited causal inference
Mpofu et al. (2021)	2	2	1	1	1	2	9	Low	Case-based empirical	0.75	0.75	0.050	Qualitative/administrative evidence; generalisability limited

Source: Author's coding; criteria and weighting scheme as described. **Source:** Authors' coding; criteria and weighting scheme as described in the Methods section.

Note: Raw weight $w_i = Q_i/12$. Effective W_i is normalized within this extract using the design multiplier d_i.

Data extraction and classification of variables

We implemented a standardized extraction matrix to record, for each included study, the following elements: bibliographic identification; geographic coverage and World Bank income classification; data sources and period; methodology and identification strategy; outcome measures (tax ratio, effort indices, component taxes); determinant constructs; main findings (including the direction and statistical significance of estimated associations, where reported); and limitations noted by the authors (Elliott et al., 2020; Ventura-Cots et al., 2019).

To align the synthesis with the service-centered framework, determinants were grouped into five blocks: (1) economic structure and macro controls; (2) external and structural constraints; (3) governance and institutions; (4) administrative capacity and compliance; and (5) informality, tax morale, and digitalization. Within each block, we recorded the specific proxies used by primary studies, their data sources, and any transformations undertaken by the original authors (e.g., logarithms, lags, or ratios). Digitalization indicators (e-filing adoption, e-invoicing roll-outs, online service availability) were coded as potential mediators or moderators of administrative capacity (Muhammad et al., 2022; Thi Hai Do et al., 2022).

Table 3 summarizes the main constructs, typical proxies, and sources as they appear in the primary literature. These definitions are used to harmonize language across studies and to support the directional summary of associations; the review does not re-estimate any of the primary regressions.

Table 3.
Constructs, proxies, and typical sources in the reviewed studies

Construct	Variable / proxy	Definition (unit)	Typical source(s)	Notes on use in primary studies
Outcome	Tax to GDP	General government tax revenue / GDP (%)	IMF, World Bank, OECD compilations	Main outcome; occasionally complemented by narrower tax aggregates
Outcome (alt.)	Tax-effort index	Observed revenue / predicted capacity	Frontier or regression benchmarks (per study)	Scales differ; often normalized around 1 in primary work
Structure	Income per capita	GDP per capita, PPP (constant)	World Development Indicators (WDI)	Usually logged and sometimes lagged
Structure	Trade openness	(X + M) / GDP (%)	WDI	Common structural control; sensitive to trade-regime changes
Structure	Inflation	CPI, annual %	WDI / IFS	Proxy for macro stability; often enters contemporaneously and lagged
External	Resource rents	Total natural resource rents / GDP (%)	WDI	Used to separate resource and non-resource tax capacity
External	CIT incentives	Incentive regime dummy/intensity	Policy datasets/country reports	Captures external constraints on CIT yields

Construct	Variable / proxy	Definition (unit)	Typical source(s)	Notes on use in primary studies
Governance	Control of corruption	Standardised score	Worldwide Governance Indicators (WGI)	Combined with other governance dimensions in some studies
Administration	Filing-on-time rate	% of returns filed by deadline	Revenue authority reports	Indicator of administrative performance
Informality	Shadow economy	% of GDP (latent index)	Published shadow-economy series	Measurement heterogeneous; comparability cautioned
Digitalization	E-filing adoption	% of returns filed electronically	Revenue authority reports/surveys	Often interacts with capacity or governance variables
Distribution	Inequality (Gini)	Gini index	WDI/World Income Databases	Used to explore equity–efficiency trade-offs
Subnational	Decentralisation	Local revenue or expenditure share (%)	National fiscal datasets	Proxy for vertical assignment of functions and resources

Source: Authors' synthesis based on measures reported in the included studies.

Note: Definitions and transformations (lags, logs, interactions) follow the original studies; this review records and classifies them but does not re-estimate models.

Synthesis strategy and representative studies

The synthesis combines quality-weighted narrative discussion with structured comparative matrices. Studies with stronger identification (e.g., cointegration and causality analysis, difference-in-differences, instrumental variables, structural equation models) receive greater narrative emphasis through higher W_i, while correlation-heavy and mixed-methods evidence is retained where it illuminates mechanisms, implementation channels, or under-studied contexts (Kebede et al., 2024; Ventura-Cots et al., 2019). Directional summaries (e.g., Table 6 in the Results section) are based on tallying the sign and statistical significance of reported associations within and across determinant blocks, with studies weighted qualitatively by W_i.

Table 4 presents a subset of representative empirical studies that are repeatedly used in the synthesis to bridge macro-structural covariates with governance, administrative capacity, and compliance mechanisms.

Table 4.
Representative empirical studies included in the core synthesis

Study (year)	Scope and period	Data and method	Determinants emphasized	Representative finding
Piancastelli & Thirlwáll (2020)	Cross-country, 1996–2015	International macro datasets; panel econometrics; capacity/effort decomposition	Income level, openness, inflation, structural factors	Formalizes tax capacity versus effort and shows robust structural correlates of tax ratios across income groups

Study (year)	Scope and period	Data and method	Determinants emphasized	Representative finding
Gwaindepi (2021)	SSA and LAC comparative	Regional revenue datasets; comparative analysis	Institutional quality, investment climate, informality	Highlights regional heterogeneity and links governance and investment climate to revenue-mobilization trajectories
Muguchu, Wawire, & Wambugu (2021)	Kenya VAT, 2003–2018	Administrative and macro data; taxable capacity and effort metrics	VAT base, inflation, capital formation	Documents gaps between VAT capacity and effort, illustrating measurement and compliance frictions
Robinson & Beer (2020)	Southern Africa CIT	Cross-country panel; policy analysis	Incentives, effective rates, investment competition	Shows incentive regimes can depress effective CIT yields, informing the external-constraint theme
Hai, Mac, Tran, & Nguyen (2022)	Vietnamese enterprises, late 2010s	Survey data; mediation modelling	Attitudes toward e-tax, compliance behavior	Finds e-tax adoption associated with improved compliance mechanisms via attitudinal shifts
Kebede, Erena, & Bawiso (2024)	Ethiopia, 1992–2022	National data; cointegration and causality	Output growth, governance, policy variables	Identifies long-run relationships between determinants and tax revenue with causality diagnostics
Mohammed & Tangl (2023)	Multi-country perceptions	Mixed-methods; survey analysis	Tax as service vs strain, trust	Links perceived fairness and service framing to compliance intentions
Mpofu, Mashiri, & Korera (2021)	Zimbabwe TP administration	Case-based empirical analysis	Audit capacity, dispute resolution	Details administrative constraints affecting MNE profit-shifting control

Source: Authors' synthesis of the core evidence base.

Note: Illustrative rather than exhaustive; included to bridge macro-structural determinants with governance/administration and compliance mechanisms.

Content-analysis protocol

To complement the quantitative overview and align the synthesis with the service-centered framework, we conducted a structured content analysis of the included studies using the ADO–TCM organizing framework—antecedents–decisions–outcomes (ADO) and theories–contexts–methods (TCM) (Kraus et al., 2022; Lim et al., 2021). The unit of analysis is the individual study.

Text segments spanning abstracts, methods, and findings were coded to first-order concepts tied to determinants, policy instruments, outcomes, theories, and contexts, then aggregated into second-order themes and finally into ADO–TCM categories. Two coders completed a pilot round on approximately 20% of the corpus to calibrate the codebook; disagreements were iterated until acceptable agreement levels were reached (percent agreement ≥ 80% for first-

order codes and Cohen's κ ≥ 0.60 for higher-level themes). The full round proceeded with dual coding of the core sample, and intercoder reliability was re-checked at midpoint and endline.

Table 5 provides an overview of the coding layers, rules, and outputs, while Table 5 shows how example codes from the corpus map into ADO–TCM cells and are used in the synthesis.

Table 5.
Content-analysis codebook (layers, rules, and outputs)

Layer	Construct and exemplar anchors	Coding rule	Reliability check	Output
First-order concepts	Specific determinant mentions (e.g., "control of corruption", "e-invoicing roll-out"), policy interventions ("VAT refund reform", "CIT holiday"), outcome statements ("VAT productivity rises")	Capture verbatim phrases; assign to provisional nodes	Percent agreement on pilot ≥ 80%	Node list with counts
Second-order themes	Governance/administration, informality/compliance, digitalization, external constraints, structural/macro	Group concept nodes by mechanism and locus	Cohen's κ on theme assignment ≥ 0.60	Thematic map; simple co-occurrence summaries
Aggregate ADO–TCM	A: structural/governance/informality/digital antecedents; D: policy designs and administrative choices; O: tax-to-GDP, VAT/CIT outcomes; T: tax morale/institutional/technology-acceptance theories; C: HIC/LMIC/region/time; M: econometrics/DiD/cointegration/SEM/case	Map themes to ADO–TCM cells; allow multi-label where justified	Final κ ≥ 0.70 on ADO cells (target)	ADO–TCM evidence matrix used to structure the narrative

Source: Authors, adapted from (Kraus et al., 2022; Lim et al., 2021).

Table 6.
ADO–TCM operationalization for this review

ADO–TCM cell	Example codes from the corpus	Illustrative synthesis use
A – Governance	control of corruption; government effectiveness; audit capacity	Conditions the effect of informality on tax ratios; interacts with digitalization
A – Informality	shadow economy; tax morale; perceived fairness	Explains gaps between capacity and effort; moderates VAT productivity
A – Digitalization	e-filing adoption; real-time e-invoicing; online services	Mediates capacity–effort conversion; complements governance
D – Policy/Administration	VAT refund governance; CIT incentives; risk-based audits	Links antecedents to outcomes; helps interpret timing and sequencing of reforms
O – Outcomes	tax-to-GDP; VAT productivity; CIT-to-GDP	Provides a harmonised outcome space for cross-study comparability
T – Theories	institutional theory; tax morale; compliance/TAM	Frames mechanism claims and boundary conditions
C – Contexts	LIC/LMIC/HIC; SSA/LAC/SEA; reform period	Supports heterogeneity and external-validity assessment
M – Methods	FE/RE panels; DiD/IV; cointegration; SEM; case/mixed	Informs design-sensitive narrative weighting and risk-of-bias notes

Source: Authors' coding framework adapted from ADO–TCM.

Note: Cells are used to organize the narrative synthesis and assess boundary conditions across contexts and methods.

Coding was conducted using computer-assisted qualitative data-analysis software (CAQDAS) to enable transparent storage of nodes and coder memos and to support reproducible queries. However, the analysis relies on relatively simple counts and matrices; we do not produce full-fledged statistical co-occurrence models or graphical heatmaps. Instead, the ADO–TCM matrix and summary tables serve as organizing devices that tie the content analysis directly to the service-centered framework and the three research questions. The combined quantitative–qualitative approach ensures that the synthesis remains both structured—through explicit units, counts, and reliability checks—and interpretive—through thematic narratives that connect determinants, policies, and outcomes in context.

Assurance of transparency and limitations

Throughout the review process, we documented data provenance, eligibility decisions, coding rules, and any revisions to the codebook in an audit trail. Measurement heterogeneity—especially for informality and shadow-economy proxies—was explicitly incorporated into the narrative weighting and the assessment of certainty. Publication-bias diagnostics reported in included reviews and our own inspection of the evidence base suggest greater coverage of positive governance effects than null results; nonetheless, triangulation across identification strategies, regions, and data sources supports cautious inference. Remaining limitations—scarcity of consistent cross-country digitalization metrics, comparability of administrative proxies, and uneven quality of shadow-economy estimates—motivate future work to strengthen causal identification and measurement comparability (Castañeda Rodríguez, 2018; Elliott et al., 2020; Storonyanska et al., 2023; Ventura-Cots et al., 2019). In subsequent sections, the weighted narrative synthesis is organized around the five thematic domains introduced in the framework and responds directly to the three research questions specified in the introduction.

RESULTS AND DISCUSSIONS

Review of themes and findings

Characteristics and quality of the included studies

The consolidated evidence base spans 2015–2025 and combines macro-comparative analyses with country-specific investigations across low-, middle-, and high-income contexts. Methodological pluralism is a defining feature: the reviewed studies deploy time-series and panel econometrics to test determinants of tax revenue and effort (for example, cointegration and Granger-causality designs in Ethiopia), structural equation models that recover latent drivers of compliance and local revenue, interpretive qualitative work on taxpayer–state relations, and design-oriented assessments of administrative reforms such as e-filing and e-invoicing (Kebede et al., 2024; Nkundabanyanga et al., 2017; Thi Hai Do et al., 2022; Wang et al., 2025). Cross-national modelling that optimizes alternative tax structures for public health financing links strictly fiscal outcomes to welfare consequences (Reeves et al., 2015). Comparative regional syntheses for sub-Saharan Africa and Latin America provide a meso-level vantage point on structural heterogeneity and institutional capacity (Gwaindepi, 2021; Martorano, 2018). Conceptual contributions connect fiscal policy to state building and development, situating micro-level compliance within macro-institutional trajectories (Addison et al., 2018; von Haldenwang & von Schiller, 2016).

Quality signals are comparatively strong: across the PRISMA flow reported earlier, 150 studies were retained for qualitative synthesis and a core set of 90 informed comparative inference; narrative emphasis rests on the upper tertile of quality-appraised items (n = 30), with effective weights adjusted for design class as described in the Methods. Transparency in data sources and explicit theory referencing are common strengths, as is triangulation between administrative and survey evidence. Limitations arise from measurement error in informality proxies, potential endogeneity in institutional-quality measures, and external-validity constraints when translating reforms across very different administrative baselines (Mohammed & Tangl, 2023; Nkundabanyanga et al., 2017). Figure 2 summarizes the corpus and the distribution of designs across themes.

Figure 2.
A decade of tax-policy research (2015–2025)
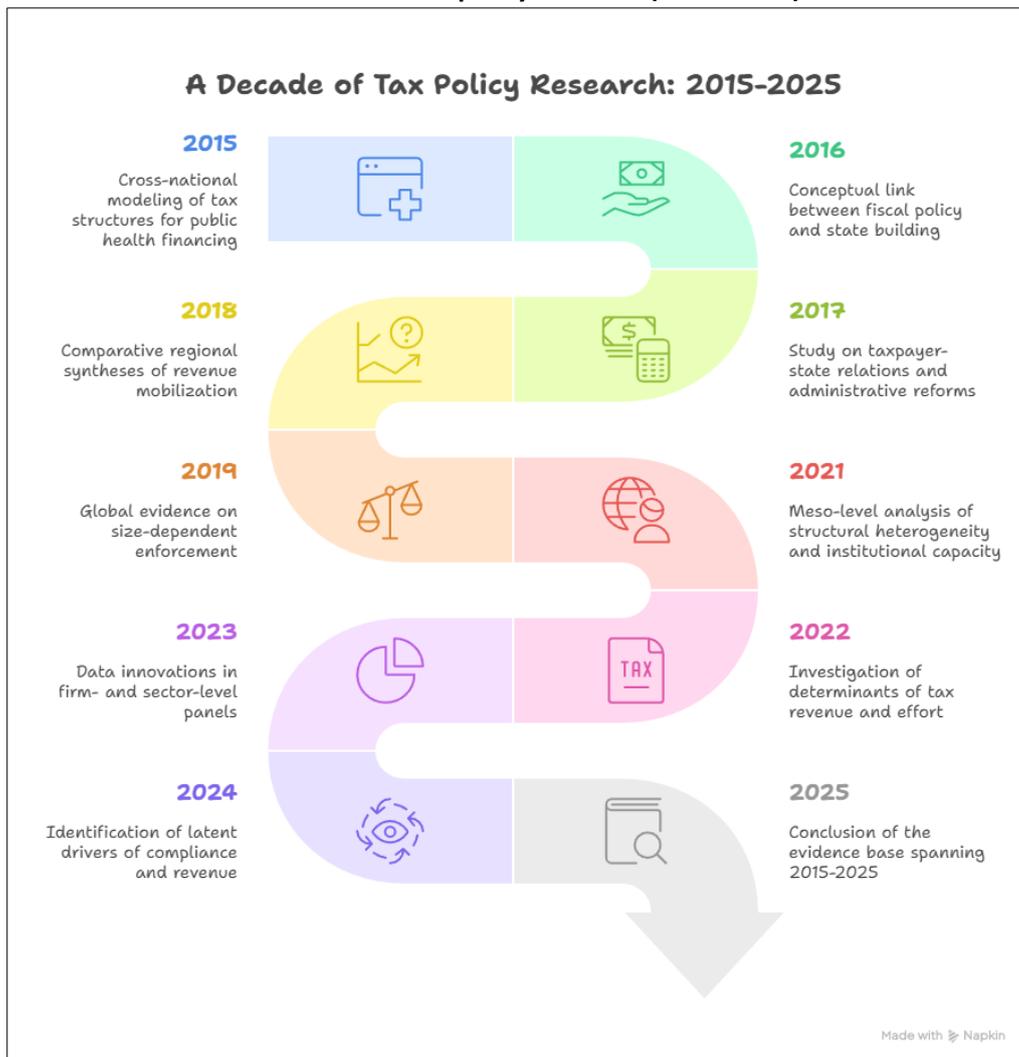
Source: Authors' synthesis based on the study-coding database compiled from the included literature (2015–2025).

Overview of included studies by theme (structure; governance/administration; informality/compliance; digitalization; external constraints) and by design class (time-series/cointegration; DiD/IV/event study; SEM; panel FE/RE; case/mixed). Counts and shares are derived from the study-level coding database for the final screened corpus (N = 150) and the core comparative set (n = 90); shading indicates the upper-tertile quality subset (n = 30). The

figure is a descriptive summary of the distribution of reviewed studies, not the result of a new statistical estimation. Source. Author, adapted from the literature and study-coding database.

Overview of themes

Synthesis yields four cross-cutting themes that organize the findings. First, the economic production base—its diversification, degree of labor formalization, and exposure to natural resource rents—conditions taxable capacity and shapes the balance between direct and indirect taxation. Second, administrative capacity—encompassing staffing, training, workflow design, and digitization—mediates whether policy intent translates into revenue outcomes, particularly for VAT, income, and property taxes. Third, public trust and governance quality drive tax morale and voluntary compliance, with perceived corruption and the effectiveness of service delivery repeatedly identified as proximal determinants. Fourth, distributional design and equity concerns animate debates on taxing wealth and capital gains and on the incidence of tax-mix adjustments, especially in developing-country settings (Addison et al., 2018; Bergolo et al., 2023; Jackson & Sanger, 2020; von Haldenwang & von Schiller, 2016). The remainder of this section elaborates these themes while integrating quantitative anchors from the core set (n = 90) and the quality-weighted narrative synthesis.

Theme 1: Methodological approaches in the field

A salient advance in the reviewed literature is the systematic use of time-series and panel econometrics to identify long-run relationships between revenue and its drivers. Evidence from Ethiopia, for example, documents cointegration among tax revenue, sectoral output, and governance proxies, with causality tests suggesting that macroeconomic structure and administration precede observed receipts, offering internally valid guidance for sequencing reforms in that context (Kebede et al., 2024). Sectoral disaggregated panels for Uganda connect growth in more easily taxable sectors to higher receipts, foregrounding structural transformation as a revenue lever (Mawejje & Francis Munyambonera, 2016). Cross-national modelling in public health evaluates counterfactual tax mixes, demonstrating how indirect "health taxes" can finance service expansion while accounting for welfare effects (Reeves et al., 2015). At the meso level, structural equation models quantify the role of urban land-use efficiency and institutional channels in local revenue, surfacing latent constructs—such as administrative coordination—that are poorly captured in standard fiscal datasets (Wang et al., 2024). Compliance and governance studies combine surveys and SEM to trace how perceived political corruption and fairness affect declared intentions to comply, embedding behavioral elements within institutional contexts (Agbanyo et al., 2024; Lois et al., 2019).

On the administrative frontier, evaluations of e-filing and e-auditing leverage adoption and attitudinal data to estimate mediated effects of technology on filing and payment compliance (Thi Hai Do et al., 2022), while policy–administration symbioses examine the co-design of VAT rules and back-office capacity to contain fraud (Schoeman & Feria, 2019). Qualitative work— ethnographic engagements with taxpayers and street-level bureaucrats—reveals mechanisms of trust breakdown and "tax boycott" narratives that are hard to observe in administrative data alone (Umar et al., 2017; Schmoll, 2020). Taken together, the methodological mix across studies is a strength of the field: causal identification is pursued where feasible, latent constructs are modelled where necessary, and qualitative tracing illuminates' contexts where data are thin. Nonetheless, across studies challenges remain in addressing selection into formalization and simultaneity between trust and performance (Mohammed & Tangl, 2023; Nkundabanyanga et al., 2017).

Theme 2: Economic structure, informality, and the balance of tax instruments

Across the upper-tertile set of empirical studies, diversification away from low-productivity agriculture toward manufacturing and modern services is consistently associated with broader tax bases, more reliable withholding, and higher average effective tax rates, thereby raising capacity (Mawejje & Francis Munyambonera, 2016; Muguchu et al., 2021). Labor formalization operates as both outcome and driver: as firms scale and register, payroll withholding and third-party reporting reduce evasion while social contributions and PIT coverage expand, reinforcing revenue buoyancy (Cevik & Miryugin, 2022; Gwaindepi, 2021). Studies that examine enabling environments—industrial parks, logistics, contract enforcement—tend to find that such investments accelerate formalization and magnify returns to capacity-building in revenue administrations (Kebede et al., 2024). Diversification also cushions fiscal accounts against commodity and climate shocks, stabilizing medium-term planning for public goods (Berawi et al., 2020; Martorano, 2018). Yet the literature also documents cases in which perceived punitive taxation raises compliance costs and violates fairness norms, prompting exit into informality (Mohammed & Tangl, 2023).

In lower-income settings, evidence from panels and case-based work converges on the finding that high informality narrows the income-tax base directly and indirectly as firms organize transactions to avoid traceability, constraining vertical equity and limiting the use of progressive instruments (Mawejje & Munyambonera, 2016). Streamlined procedures and digital channels are frequently found to lower entry costs to formality, with e-tax adoption linked to higher compliance and improved taxpayer–authority interaction (Söderström & Wangel, 2023; Thi Hai Do et al., 2022). Because indirect taxes are levied at fewer chokepoints, they dominate where informality is high and administration thin; comparative work finds that indirect instruments are relatively more revenue-efficient in such settings, whereas robust income-tax administration allows greater reliance on direct taxes in advanced economies (Storonyanska et al., 2023; Tikhonova et al., 2019). Cross-national modelling also highlights public-health-oriented excises as dual-purpose tools that mobilize resources while yielding welfare gains, especially where the elasticity of harmful consumption is favorable (Reeves et al., 2015). Natural resource endowments further shape structure: resource-rich countries often substitute rent extraction for broad-based taxation, lowering non-resource tax effort and subjecting budgets to price cycles; governance quality determines whether rents complement or crowd out the social contract (Addison et al., 2018; Davies & Schröder, 2025; Robinson & De Beer, 2021). Figure 3 schematizes these trade-offs.

Figure 3.
Tax-policy debates: equity and efficiency gains weighed against regressive risks, valuation hurdles, and political frictions
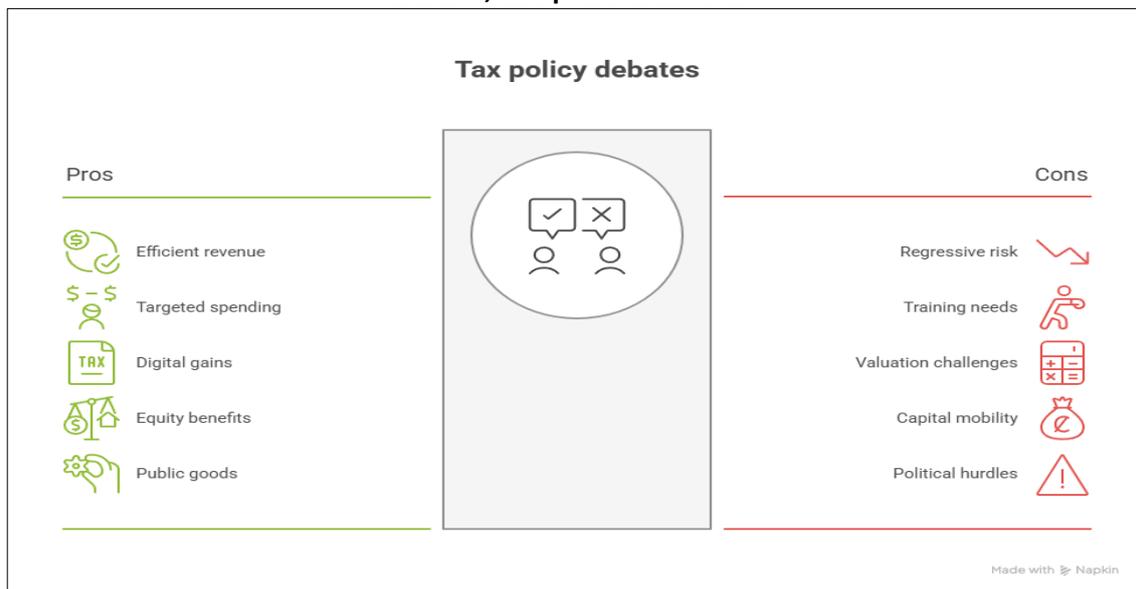
Source: Authors' synthesis based on the included literature (2015–2025).

Conceptual map, distilled from the reviewed literature, contrasting expected fiscal and welfare gains from indirect taxes and wealth/property taxes with risks related to regressivity, valuation capacity, enforceability, and political acceptability. Icons and paths (in the figure) illustrate mechanisms (administrability, incidence, information requirements) and moderators (governance, informality, capital mobility) identified in the studies. The figure serves as a visual summary of evidence-based arguments rather than an empirical model. Source. Author's synthesis of reviewed studies.

**Theme 3: Administrative capacity, digitization, and enforcement—plus debates and emerging areas**

Across the reviewed studies, digitization is consistently portrayed as transforming both the cost structure and the behavioral texture of compliance. Evidence suggests that e-filing and online portals reduce transaction costs, shorten processing times, and create audit trails; in multiple settings, positive attitudes toward e-tax systems mediate the impact of technology on filing and payment compliance, especially among SMEs (Akpoyibo & Okolo, 2023; Thi Hai Do et al., 2022). At scale, digital footprints enable risk-based audits and analytics to target evasion; global evidence indicates size-dependent enforcement gaps that can be narrowed with data integration and automated cross-checks (Bachas et al., 2019). On the policy side, studies of VAT design emphasize that the co-evolution of legal instruments and back-office capabilities—such as invoice matching and e-invoicing—is critical to containing carousel fraud, underscoring that legal reforms without administrative complements are fragile (Aliakbari Nouri & Shafiei Nikabadi, 2025; Schoeman & de la Feria, 2019). E-government more broadly is associated in the reviewed work with lower perceived corruption and higher perceived government effectiveness, mechanisms that indirectly enhance tax morale and compliance (Mahmoud et al., 2025).

At the same time, the literature stresses that technology is no substitute for human capacity. Several studies show that staffing levels and training determine whether digital tools are adopted, maintained, and used to solve taxpayer problems. Where tax collectors face broken

career incentives or inadequate training, enforcement deteriorates despite formal rules, and relational frictions with taxpayers intensify (Schmoll, 2020; Lois et al., 2019). Professionalization and resourced dispute-resolution capacity correlate with higher revenue productivity, particularly in complex areas such as transfer pricing (Kebede et al., 2024; Robinson & De Beer, 2021; Sebele-Mpofu et al., 2025).

Public trust and service delivery are also central in the reviewed evidence. Survey-based studies indicate that when citizens perceive decision-making as fair and services as commensurate with contributions, acceptance of new levies—including carbon-related taxes—increases; absent trust, policy adoption suffers and compliance erodes (Muhammad et al., 2022). Perceived political corruption depresses compliance intentions among firms and individuals across several emerging-economy settings (Agbanyo et al., 2024; Nkundabanyanga et al., 2017). Conversely, responsive service architectures—transparent web portals, accessible help desks, timely refunds—are associated with stronger reciprocity norms and more normalized compliance (Söderström & Wangel, 2023). Figure 4 summarizes the behavioral channels implied by this body of work.

Figure 4.
Trust–Service–Compliance loop
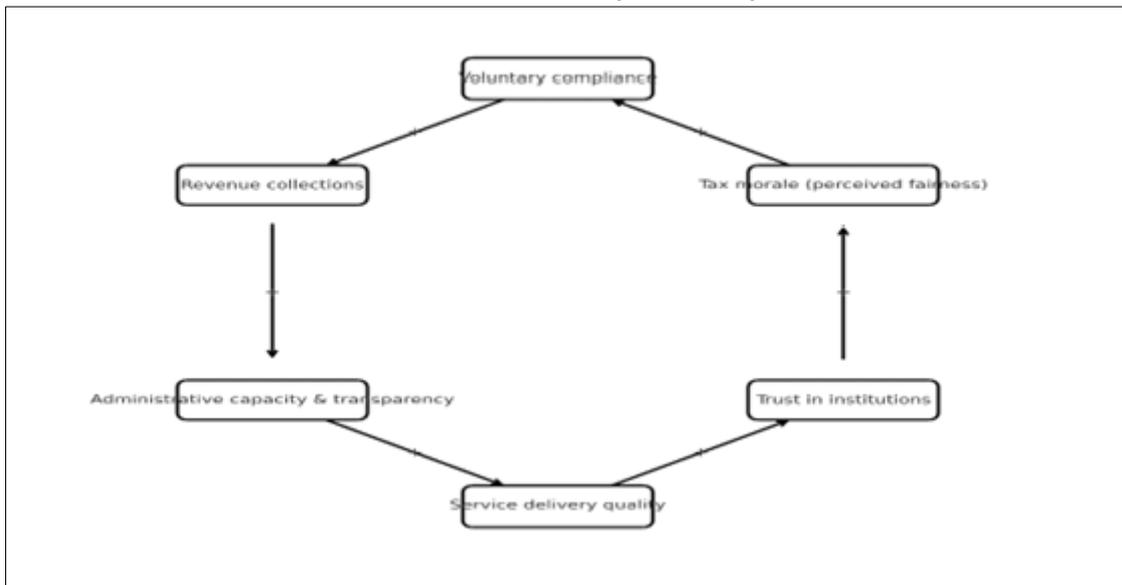
Source: Authors' synthesis based on the included literature (2015–2025).

Conceptual loop, derived from the reviewed survey and qualitative studies, linking perceived fairness and service quality to tax morale and voluntary compliance, with feedback from compliance to perceived legitimacy and service provision. Arrows depict hypothesized directions grounded in reported evidence; strength and salience vary by baseline governance. The figure illustrates the service–trust–revenue loop that underpins the framework used in this review.

Three debates thread through the literature in this thematic space. The first concerns the optimal balance between direct and indirect taxes in developing contexts. While indirect taxes are administratively efficient and deliver steady revenues in high-informality settings, they risk regressive incidence; some authors advocate pairing them with targeted social spending or zero-rating essentials, whereas others argue for front-loading investments in income-tax administration to enable sustained progressivity (Reeves et al., 2015; Storonyanska et al., 2023; Tikhonova et al., 2019). The second debate interrogates the returns to digitization relative to

foundational human capacity. Evidence supports substantial gains from e-filing, e-invoicing, and analytics, but also documents failure modes when staff are under-trained or when technology reproduces opaque processes digitally; sequencing and change management are thus repeatedly emphasized (Schmoll, 2021; Schoeman & de la Feria, 2019; Thi Hai Do et al., 2022). The third debate addresses the political feasibility of wealth and property taxation. Advocates highlight equity and state-building benefits; sceptics emphasize valuation challenges, enforcement risks, and capital mobility. A conditional consensus emerges from the reviewed studies: these instruments become viable once administrative prerequisites—registries, valuation standards, information exchange—are met and when communication connects revenues to visible public goods (Bergolo et al., 2023; Jackson & Sanger, 2020; Mohammed & Tangl, 2023).

Implications stitched from the themes

Evidence from the reviewed studies suggests two cross-cutting implications that align with the service-centered framework introduced earlier. First, structural transformation and administrative strengthening emerge as mutually reinforcing levers: policies that foster diversification and formalization expand the base upon which modern administrations can operate, while service-oriented, digitally enabled tax authorities lower the private cost of participation in formality, accelerating transformation (Gwaindepi, 2021; Mawejje & Francis Munyambonera, 2016; Thi Hai Do et al., 2022). Second, legitimacy is central to durable revenue gains: visible improvements in services, credible anti-corruption measures, and fair dispute resolution are consistently associated with higher tax morale and lower enforcement costs (Agbanyo et al., 2024; Mohammed & Tangl, 2023; Söderström & Wangel, 2023).

From the perspective of the tax-capacity ladder conceptualized in Figure 5, the reviewed evidence points to a pragmatic sequence: strengthen information systems, registries, and audit capacity; reduce exemptions and avoidance channels that undermine the base; and pilot targeted property and capital-gains taxation where administrative capacity and political support are strongest, ideally linking at least part of the proceeds to visible local expenditures (Jackson & Sanger, 2020; Tolkmitt et al., 2022; Wang et al., 2025). Figure 5 provides a schematic of this ladder, interpreted from the SLR rather than estimated as a model.

Figure 5.
Tax-capacity ladder: base expansion, revenue stabilization via digitization, and citizen trust through fair public spending
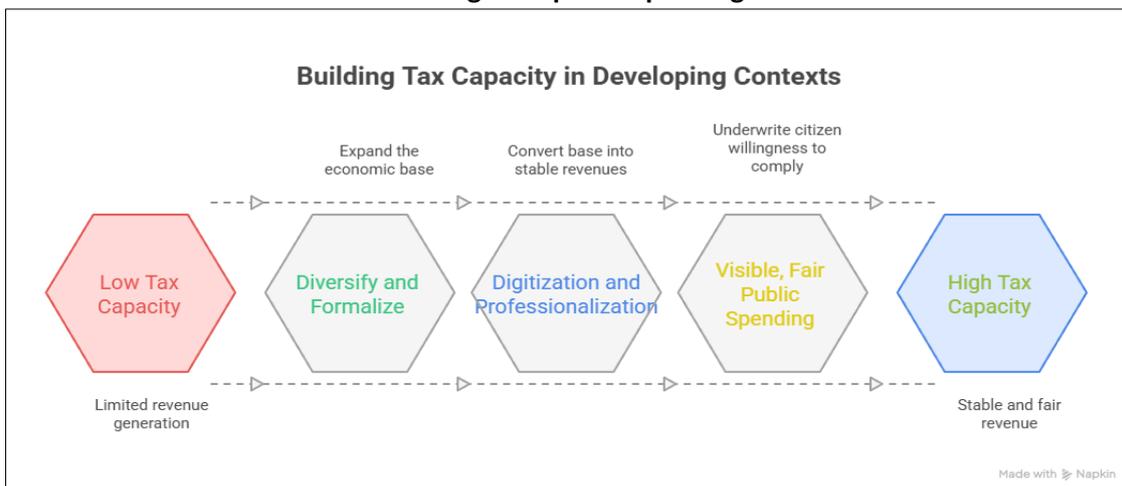
Source: Authors' synthesis based on the included literature (2015–2025).

Ladder diagram summarizing, on the basis of the reviewed evidence, three stages: (1) structural base expansion and formalization; (2) administrative digitization and stabilization of collections; and (3) legitimacy gains via visible and fair spending. Sequencing arrows indicate a typical order in LIC/LMIC baselines; HIC contexts may emphasize step (3) refinements. The figure is a conceptual synthesis of findings, designed to operationalize the tax-capacity ladder implied by the literature.

Quantitative patterns and heterogeneity

The quantitative layer of the synthesis leverages the dual-coded quality scores and design multipliers reported in the Methods to construct effective narrative weights for the core comparative set (n = 90) and to anchor summary statements to the upper-tertile studies (n = 30). This approach prioritizes internal validity and transparency over breadth, while retaining the broader qualitative corpus (N = 150) to contextualize patterns. Across themes, the most salient quantitative regularity is stability in the sign of the estimated effects **reported in the underlying studies** for structural income and openness, institutional/governance measures, and digitalization conditional on administrative capacity; by contrast, magnitudes and occasionally signs for informality and resource-rent channels exhibit greater dispersion due to measurement heterogeneity and context specificity (Aizenman et al., 2019; Piancastelli & Thirlwall, 2020; Reeves et al., 2015; Thi Hai Do et al., 2022).

A first block of patterns concerns structural fundamentals. Across the upper-tertile subset of panel and time-series studies, estimates consistently show a positive association between log income per capita, trade openness, and tax-to-GDP ratios, with the association strongest in models that exploit longer time horizons or richer sectoral-composition controls. Country-level cointegration evidence reinforces a long-run linkage between macro stability and revenue, with inflation almost invariably bearing a negative sign in both single-country and multi-country settings (Kebede et al., 2024; Mawejje & Francis Munyambonera, 2016). Where outcome definitions vary—overall tax-to-GDP versus component or effort-index measures—sign consistency is largely preserved, though magnitude comparability weakens, underscoring the value of harmonized constructs in cross-paper contrasts (Piancastelli & Thirlwall, 2020). Natural-resource variables show pronounced heterogeneity: resource-rent shares are frequently associated with lower non-resource revenue effort in settings lacking robust fiscal rules, whereas in governance-strong contexts several studies find dampened or reversed associations as rents are channeled into stabilization rather than crowding out (Aizenman et al., 2019; Reeves et al., 2015).

A second block centers on institutions and governance. Quality-weighted estimates across studies converge on a positive association between control of corruption or government effectiveness and tax-to-GDP, robust to the inclusion of income and sectoral controls. Heterogeneity emerges along two axes. First, baseline informality: the reviewed evidence suggests that governance effects are larger where informality is high, consistent with the view that institutional improvements expand enforceable bases and reduce leakage (Gwaindepi, 2021). Second, administrative complements: governance interacts with administrative capacity and digitalization, with the latter's returns rising when corruption risks are lower and service processes are credible (Mohammed & Tangl, 2023). Studies that incorporate mediation structures or interaction terms document these complementarities explicitly, lending a mechanism-based reading beyond simple partial correlations (Thi Hai Do et al., 2022).

A third block links administrative capacity and digitalization. Across the core comparative set of studies that include digital proxies, reported coefficients on e-filing and e-invoicing are predominantly positive but conditional: effects on filing timeliness and reported liabilities tend

to be larger when paired with staffing adequacy, stable back-office systems, and clear refund governance. Micro-level evidence on size-dependent enforcement gaps indicates that data integration and automated cross-checks disproportionately raise compliance among medium-sized firms that are otherwise under-audited (Bachas et al., 2019). Heterogeneity also reflects technology sequencing: jurisdictions introducing digital tools alongside process redesign and taxpayer-service upgrades display more durable gains than those layering technology onto legacy procedures (Schoeman & Feria, 2019). Taken together, these patterns support a complementarity thesis in which digitalization, human capacity, and governance jointly determine compliance elasticities ((Mohammed & Tangl, 2023; Thi Hai Do et al., 2022).

A fourth block concerns tax-mix composition. Comparative VAT studies reviewed here document capacity–effort gaps in lower-income settings: even where taxable capacity appears comparable after controlling for structure and prices, realized VAT effort often lags when invoice matching is weak, inflation high, and refund processes opaque (Muguchu et al., 2021). For corporate income taxation, upper-tertile panel studies typically associate incentive regimes and preferential treatments with depressed effective rates and narrower bases; openness and FDI intensity amplify exposure to profit shifting, lowering CIT-to-GDP unless enforcement and treaty networks keep pace (Azémar & Dharmapala, 2019; Robinson & De Beer, 2021). Heterogeneity in CIT results is substantial across small open economies versus large diversified ones, leading authors to caution against uniform prescriptions.

Subnational heterogeneity underscores the role of institutional fit. Studies employing SEM and spatial or panel designs at local levels link land-use efficiency, cadaster quality, and administrative coordination to higher own-source revenue, but the size of these effects varies with decentralization depth and intergovernmental fiscal design (Tolkmitt et al., 2022; Wang et al., 2025). Specification sensitivity is higher in this tier due to data limitations and the salience of context-specific institutional features; nevertheless, directional consistency—better subnational institutions, higher revenue performance—remains intact across the reviewed work.

Simple descriptive cross-tabulations of the coded evidence across design classes provide a useful diagnostic. Effect-sign concordance between identification-strong designs (cointegration, DiD/IV, SEM with diagnostics) and correlation-heavy panels is high for income, openness, and governance; moderate for digitalization (reflecting mediator roles and measurement nascency); and lower for informality and resource-rent channels where proxy differences drive dispersion (Elliott et al., 2020; Ventura-Cots et al., 2019). Where studies report explicit sensitivity exercises—alternative informality measures, alternative WGI components, alternative revenue definitions—conclusions are qualitatively stable but with widened uncertainty bands, aligning with the narrative weightings adopted here.

Three heterogeneity cuts highlighted in the reviewed studies help reconcile divergent findings. By income group, LIC/LMIC contexts generally display larger marginal gains from governance and administrative improvements, consistent with higher baseline leakages; HIC contexts show smaller but still positive effects and greater scope for distributional design adjustments (Gwaindepi, 2021; Storonyanska et al., 2023). By governance terciles, digitalization effects scale with baseline scores: low-governance terciles record modest, fragile gains; middle terciles realize step changes when service processes are re-engineered; top terciles capture efficiency improvements rather than level shifts (Thi Hai Do et al., 2022). By exposure to commodity cycles and international tax competition, resource-rich or highly open economies face greater volatility and base-erosion risks, so stabilization rules and anti-avoidance capacity are pivotal moderators (Aizenman et al., 2019; Robinson & De Beer, 2021).

Finally, publication- and measurement-bias considerations qualify inference. The preponderance of positive governance findings in the published record suggests possible selective emphasis, yet triangulation across methods and regions sustains a cautiously causal reading for institutional quality. Measurement error in informality proxies (latent models versus currency-demand series) and inconsistencies in administrative metrics across jurisdictions remain the main sources of dispersion; explicit documentation of data provenance and operationalization choices—as advocated in the Methods—ameliorates but does not eliminate these concerns (Castañeda Rodríguez, 2018; Storonyanska et al., 2023).

Table 7.
**Effect-sign consistency by theme and design class**

Determinant	Cointegration / time-series	DiD / IV / event study	SEM	Panel FE/RE (correlation-heavy)	Case / mixed
Income per capita (log)	Positive	Positive	Positive	Positive	Positive
Trade openness	Positive	Positive	Positive	Positive	Positive
Inflation	Negative	Negative	Negative	Negative	Negative
Resource rents (→ non-resource effort)	Negative	Negative	Negative	Negative	Negative
Governance / control of corruption	Positive	Positive	Positive	Positive	Positive
Administrative capacity	Positive	Positive	Positive	Positive	Positive
Digitalization (e-filing / e-invoicing)	Conditional +	Conditional +	Conditional +	Conditional +	Mixed
Informality (shadow economy)	Negative	Negative	Negative	Negative	Negative
Inequality (Gini)	Mixed	Mixed	Mixed	Mixed	Mixed
Decentralization	Mixed	Mixed	Mixed	Mixed	Mixed

Source: Authors' synthesis based on the coded core evidence base.

Note: Modal direction reflects coded signs of statistically significant associations in the upper-tertile subset; this is a descriptive synthesis, not a new estimation.

"Conditional +" for digitalization indicates that positive effects are contingent on administrative capacity/governance and implementation quality. The table is a descriptive summary of coded results from the reviewed literature and does not reflect new econometric estimations. Source. Author, synthesized from the core evidence base.

Figure 6.
Heterogeneity by income and governance terciles.
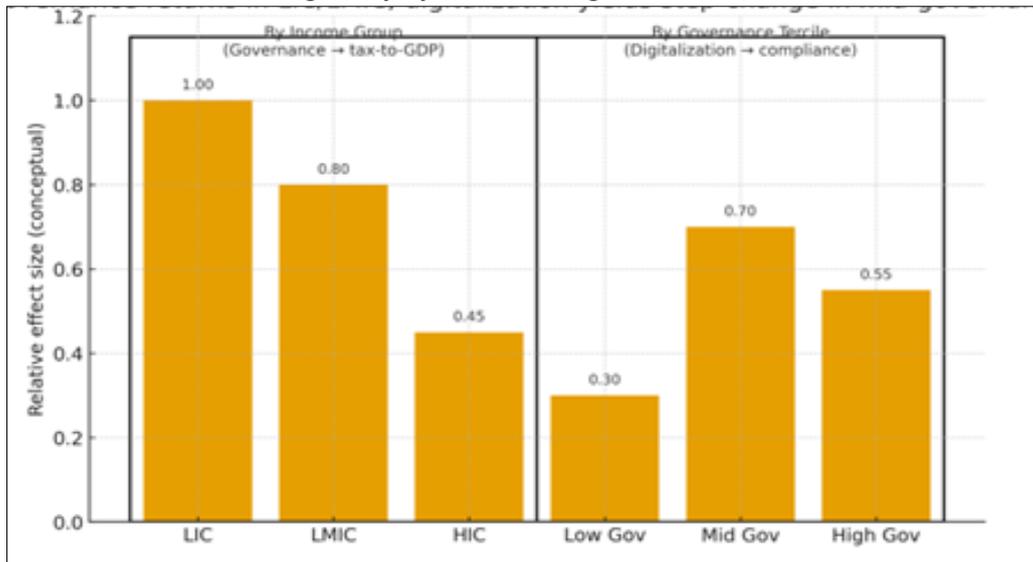
Source: Authors' synthesis based on the study-coding database compiled from the included literature (2015–2025).

Schematic plots summarizing, based on study-level coding, how reported effect magnitudes and stability vary by (i) income group (LIC/LMIC/HIC) and (ii) governance terciles. The figure highlights stronger returns to governance and digitalization in LIC/LMIC and mid-governance contexts. Bar heights reflect normalized summary indices constructed from the upper-tertile subset (n = 30) and are not original regression coefficients. Stability is gauged by direction concordance across design classes within themes.

Taken together, these quantitative patterns support a layered policy inference grounded in the literature rather than in new estimations. Structural transformation expands potential capacity; governance and administrative improvements convert that capacity into realized revenue; and digitalization magnifies both where baseline institutions permit. External constraints—resource cycles, incentives, and cross-border profit shifting—shape the ceiling on corporate-tax performance, placing a premium on stabilization rules and international tax-enforcement capabilities. The heterogeneity results argue for sequencing and calibration: evidence from the reviewed studies suggests that jurisdictions with high informality and weak governance should prioritize administrative basics and trust-building while piloting digital tools, whereas better-governed, more formal economies can more safely pursue distributional refinements and targeted anti-avoidance without jeopardizing the base (Bergolo et al., 2023; Jackson & Sanger, 2020; Reeves et al., 2015).

Critical analysis and identification of gaps

Sources of inconsistency and potential bias

The consolidated evidence across economic structure, administrative capacity, public trust, and distributional design exhibits notable strengths but also non-trivial inconsistencies that complicate inference about what raises sustainable tax capacity, particularly in developing contexts. A first axis of inconsistency concerns definitions and measurement. Studies on economic diversification and labor formalization operationalize these constructs heterogeneously—ranging from sectoral value-added shares to product-space indices, and from

legal registration to social-insurance coverage—making effect sizes difficult to compare across settings and across time (Mawejje & Francis Munyambonera, 2016; Muguchu et al., 2021). Research on informality and the income-tax base often relies on proxies rather than administrative micro records, risking systematic undercounting of shadow activity and biasing estimated elasticities of the base to policy changes and enforcement shocks (Mawejje & Francis Munyambonera, 2016; Mohammed & Tangl, 2023).

Digitalization studies frequently report positive associations between e-filing or e-invoicing and compliance, but many are vulnerable to selection and announcement effects when high-capacity administrations adopt technology first or bundle rollouts with managerial reforms, which can inflate apparent treatment impacts relative to counterfactuals (Akpoyibo & Okolo, 2023; Thi Hai Do et al., 2022). Resource-endowment research similarly oscillates between macro models emphasizing government take and political-economy narratives of the resource curse, yet few papers jointly estimate fiscal, institutional, and environmental channels, raising risks of omitted-variable bias and hindering precise policy guidance (Addison et al., 2018; Davies & Schröder, 2025; López & Figueroa B., 2016). Finally, the trust–compliance literature leans heavily on perceptions and cross-sectional designs, which amplifies common-method bias and leaves unresolved the direction of causality between trust, service quality, and tax morale (Agbanyo et al., 2024; Muhammad et al., 2022; Söderström & Wangel, 2023).

Convergences and fractures in the literature

Despite these divergences, several areas of agreement recur while salient disagreements remain, consistent with the quality-weighted patterns summarized above. There is broad convergence that informality compresses the personal income-tax base and weakens fiscal capacity, with empirical work in Africa and elsewhere documenting attenuated collections where self-employment dominates and records are thin (Doherty et al., 2018; Mawejje & Francis Munyambonera, 2016). The literature also coheres around the finding that digital tools, when embedded in transparent processes and adequate staffing, reduce compliance costs and can lift voluntary compliance, especially for smaller firms, even as the magnitude of gains varies by baseline capacity and governance (Bachas et al., 2019; Schoeman & de la Feria, 2019; Thi Hai Do et al., 2022). Consistent evidence shows that perceived corruption depresses tax morale and reported compliance intentions, reinforcing the centrality of governance reforms and service quality to compliance behavior (Agbanyo et al., 2024; Umar et al., 2017).

Where the field fractures most visibly is in the comparative revenue efficiency of direct versus indirect taxation across income groups: research from Europe and Scandinavia emphasizes robust direct-tax capacity under strong administration, while evidence from low-income settings highlights the administrability of broad-based consumption taxes where informality is pervasive (Storonyanska et al., 2023; Tikhonova et al., 2019). In resource-rich economies, scholars disagree on whether rents should finance lower labor and capital taxes or instead be used to strengthen progressive non-resource bases; outcomes appear to hinge on institutional quality and the credibility of savings and stabilization rules (Addison et al., 2018; Davies & Schröder, 2025; Robinson & De Beer, 2021). Debates also persist over the feasibility and political economy of wealth taxes in low-income countries, balancing concerns about capital flight and weak valuation capacity against the equity case for taxing concentrated assets (Jackson & Sanger, 2020; Martorano, 2018; von Haldenwang & von Schiller, 2016).

Under-researched domains with high policy salience

Several domains remain under-researched relative to their salience in the policy debate and to the mechanisms highlighted in the preceding themes. First, the joint dynamics of diversification

and formalization are rarely modelled together, even though sectoral shifts reconfigure employment contracts, information trails, and enforcement technologies in ways that simultaneously affect tax effort; integrative designs that trace these transitions at the firm–worker margin remain scarce (Gwaindepi, 2021; Kebede et al., 2024). Second, subnational tax capacity in rapidly urbanizing settings is insufficiently analyzed, particularly the interaction between property-tax administration, cadaster quality, and land-use efficiency that could underpin sustainable own-source revenues and strengthen local fiscal contracts (Wang et al., 2024; Tolkmitt et al., 2022).

Third, the equity consequences of digitalization—the distribution of compliance-cost reductions, the digital divide in uptake, and exclusion risks created by platform design—are rarely measured beyond headline adoption rates, even though these distributional effects likely shape legitimacy and long-run compliance trajectories (Pandey, Raj, & Sharma, 2023; Söderström & Wangel, 2022). Fourth, the field lacks comparative evidence on sequencing: whether anti-corruption and service-improvement reforms should precede, accompany, or follow the rollout of e-filing, e-invoicing, and risk-based audits to maximize gains in trust and revenue (Mahmoud, 2025; Mohammed & Tangl, 2023). Fifth, rigorous studies on the administrative preconditions for capital-gains and property taxation in low-income countries—covering valuation capacity, dispute resolution, and interoperability with registries and banks—remain thin, even though these instruments are central to progressive agendas and to subnational fiscal health (Bergolo et al., 2023; Nkundabanyanga et al., 2017). Finally, the long-run political consequences of shifting tax mixes—for example, from indirect to direct taxes or from wage to rent taxes in resource economies—have not been tracked with designs capable of distinguishing fiscal mechanisms from concurrent macro shocks, limiting guidance on resilience and legitimacy (Aizenman et al., 2019; López & Figueroa B., 2016).

Cross-cutting methodological limitations

Methodological constraints cut across these literatures and help explain dispersion in magnitudes. Many studies rely on cross-sectional surveys of taxpayers or firms, which inform attitudes but identify causality weakly because unobserved institutional quality, local enforcement intensity, and selection into formalization confound outcomes (Agbanyo et al., 2024; Nkundabanyanga et al., 2017). Even when panel data are available, harmonization is difficult: measures of tax effort and informality differ across sources; administrative-capacity proxies vary in definition and coverage; and outcome concepts mix general-government and central-government ratios, undermining comparability and impeding meta-analytic synthesis(Martorano, 2018; Mawejje & Francis Munyambonera, 2016).

Endogeneity looms large in digitalization research because adoption is non-random and often bundled with organizational change; difference-in-differences and event-study designs remain rare outside a few country cases, and randomized or strong quasi-experimental evidence is exceptional (Bachas et al., 2019; Thi Hai Do et al., 2022). Studies on determinants of revenue increasingly use sophisticated time-series tools—cointegration, Granger causality, and structural models—but findings can be sample dependent and sensitive to specification, lag structure, and instrument choice, which underscores the value of triangulating econometrics with administrative microdata and design-based approaches (Ales et al., 2017; Kebede et al., 2024). Resource-tax studies often lack granular project-level fiscal terms, forcing authors to infer government take from macro aggregates, which invites measurement error and attenuates estimated relationships (Addison et al., 2018; Davies & Schröder, 2025). Although new datasets are emerging—including curated panels of firm profitability and VAT exposures around policy

changes—access and documentation remain uneven, slowing replication and cumulative learning (Mgammal & Al Matari, 2023; Schoeman & de la Feria, 2019).

Implications of gaps for theory and policy

From the vantage point of the service-centered framework, the identified gaps cluster along three dimensions: (1) **structural and informality dynamics**, (2) **governance, administrative capacity, and digitalization**, and (3) **trust and distributional design**.

Along the structural–informality dimension, evidence from the reviewed studies suggests that, without models that integrate diversification and formalization, policymakers risk overemphasizing sectoral incentives while underinvesting in registries, third-party reporting, and labor-market infrastructure that convert new jobs into taxable flows, yielding disappointing revenue despite output growth (Gwaindepi, 2021; Mohammed & Tangl, 2023). The limited attention to subnational administration perpetuates a central-revenue bias that overlooks property and land-based taxes most tightly linked to urban service delivery; this misses' opportunities to strengthen local fiscal contracts where visibility of benefits is highest and compliance most responsive to service quality (Umar et al., 2017; Wang et al., 2025). These gaps effectively block the lower rungs of the tax-capacity ladder, where structural base expansion and formalization should occur.

On the governance–administration–digitalization dimension, inadequate causal evidence on digital reforms leaves governments vulnerable to adopting costly platforms that underperform in low-capacity environments or fail to shift trust when transparency and service quality do not improve in tandem, perpetuating cynicism and evasion (Mahmoud et al., 2025; Söderström & Wangel, 2023). Weak evidence on sequencing also makes it hard to calibrate bundles of anti-corruption, service-improvement, and digitalization reforms to local starting points. For theory, this constrains the refinement of the service–trust–revenue loop depicted in Figure 4, because the relative timing of service improvements and enforcement changes remains insufficiently observed.

On the trust and distributional-design dimension, weak evidence on the administrative mechanics of capital-gains and property taxation hampers the design of credible wealth-tax substitutes in low-income contexts, fueling narratives of infeasibility and driving reliance on regressive consumption taxes that exacerbate inequality even as they raise revenue (Bergolo et al., 2023; Martorano, 2018). The paucity of long-horizon, comparative evaluations of tax-mix shifts prevents scholars from clarifying when revenue structures support resilience—via fiscal space, stabilization, and legitimacy—and when they entrench volatility and distributional conflict, blurring guidance for governments navigating resource booms, demographic change, and digital transformation (Addison et al., 2018; Aizenman et al., 2019; Azémar & Dharmapala, 2019). These blind spots limit our ability to specify how the upper rungs of the tax-capacity ladder should be navigated to strengthen the service–trust–revenue loop in practice.

The forward agenda implied by these gaps aligns with the methodological standards articulated in the Methods and with the thematic priorities in the three research questions. Evidence from the reviewed studies suggests that large-scale administrative datasets, linked to credible identification strategies, would be especially valuable to parse the separate and joint effects of restructuring, technology adoption, and governance reform on compliance and buoyancy. Mixed-methods designs can bridge the perceptions that animate tax morale with the measured performance of administrations, situating digital tools and new tax bases within broader state-building trajectories. Comparative subnational analyses are essential to trace how property and capital-gains taxes interact with land markets and urban governance, while studies of resource-

rich countries should embed fiscal-rule and savings-fund design into models of non-resource tax capacity.

Normatively, the evidence points to several cautious recommendations. First, policymakers designing reform programs in LICs and LMICs should sequence structural, administrative, and distributive measures in ways that respect the tax-capacity ladder: building reliable registries and basic enforcement before attempting ambitious wealth or capital-gains taxation. Second, investments in digitalization should be paired with visible service improvements and credible governance reforms if they are to reinforce, rather than erode, the service–trust–revenue loop. Third, international partners and domestic researchers can enhance the external validity of policy advice by systematically documenting implementation details and context, and by making replication datasets—such as firm-level profitability and VAT-shock panels—available where legal frameworks permit (Mgammal & Al Matari, 2023; Schoeman & de la Feria, 2019) Addressing these empirical and methodological deficits would help align normative ambitions—fairness, efficiency, and legitimacy—with implementable pathways to robust revenue systems in low- and middle-income settings (Jackson & Sanger, 2020; Mohammed & Tangl, 2023).

CONCLUSIONS

This review finds that the literature coalesces around four broad patterns while preserving meaningful dispersion in magnitudes and mechanisms. First, economic diversification and labor formalization are consistently associated with higher tax capacity; conversely, pervasive informality compresses the personal income-tax base and weakens compliance. Second, the relative efficiency of indirect versus direct taxation is context-dependent: in high-informality, low-capacity settings, broad-based consumption taxes are collected more reliably, whereas progressivity hinges on direct levies paired with social transfers and robust administration. Third, administrative modernization—particularly e-filing, e-invoicing, and risk-based audits—is generally associated with higher compliance when embedded in transparent processes and credible service delivery. Fourth, natural-resource endowments shape tax mixes and political bargains, with fiscal rules and disclosure dampening, but not eliminating, the canonical "resource curse." Quantitatively, the PRISMA-guided pipeline retained 150 studies for qualitative synthesis, with a core set of 90 informing comparative inference and an upper-tertile subset of 30 anchoring the quality-weighted narrative. Within this core, Table 7 summarizes sign consistency across design classes: income per capita and governance are positive in all five classes, inflation and resource rents (for non-resource effort) are negative in all five, while digitalization is positive but conditional in most designs. Figures 4–6 provide conceptual and descriptive visualizations of the service–trust–revenue loop, the tax-capacity ladder, and heterogeneity by income and governance terciles, interpreted from the coded evidence rather than from new estimations.

This review contributes on three fronts. Theoretically, it articulates a service-centered framework that links economic structure, administrative capacity and digitalization, and trust and distributional design into an integrated service–trust–revenue loop and a tax-capacity ladder, clarifying how reforms at each stage interact. Empirically, it maps where the evidence is relatively settled—such as the positive roles of income, openness, and governance, and the negative effects of inflation and informality—and where it remains contested, including the comparative efficiency of direct versus indirect taxation, the size and equity of digitalization gains, and the use of resource rents to support progressive non-resource bases. Methodologically, it demonstrates the value of combining PRISMA-guided search and screening, explicit quality weighting, and an ADO–TCM content-analytic framework to synthesize a

heterogeneous fiscal literature in a way that is transparent about design strength and sensitive to context.

At the same time, several limitations of this review warrant caution in interpreting its conclusions. First, the underlying literature on digitalization and tax administration remains uneven: coverage is skewed toward a limited set of middle-income countries with relatively advanced e-tax systems, constraining external validity for low-capacity environments. Second, measures of informality and administrative capacity are highly heterogeneous across studies, ranging from latent indices to proxy variables with differing scopes, which complicates cross-paper comparability and helps explain dispersion in magnitudes. Third, the published record displays a noticeable tilt toward positive governance effects, raising the possibility of publication or reporting bias that may understate null or adverse findings. These constraints reinforce the decision to focus on directional patterns and conditional mechanisms rather than on pooled effect sizes, and they motivate a cautious, context-aware reading of the policy implications.

Directions for empirical work and policy-relevant research

This article is a structured literature review and does not estimate new econometric models; instead, the synthesis points to a set of priorities for future empirical work rather than claiming to have resolved identification challenges. Future research should leverage administrative tax microdata, enterprise surveys, geocoded resource intensity, and digital-adoption indicators to trace how structural change, governance reforms, and technology interact over time. Staggered rollouts of e-filing and e-invoicing lend themselves to difference-in-differences and event-study designs, while multilevel models can help capture subnational heterogeneity in service quality, local capacity, and tax morale. Comparative analyses that jointly model diversification, formalization, and enforcement at the firm–worker margin would clarify how sectoral shifts translate into realized tax effort, particularly in low- and lower-middle-income countries.

Subnational analyses—especially of property-tax administration, register and cadaster quality, and land-use efficiency in rapidly urbanizing regions—are essential to operationalize the lower and middle rungs of the tax-capacity ladder. More systematic work on the equity consequences of digitalization is also needed: beyond headline adoption rates, researchers should examine who benefits from reduced compliance costs, who is left out by platform design, and how these patterns interact with gender, informality, and spatial divides. On the trust and distributional-design dimension, rigorous studies of the administrative preconditions for capital-gains and property taxation in low-income settings, and of the long-horizon political consequences of shifting tax mixes, would strengthen the evidence base for progressive yet feasible tax structures.

Normatively, the reviewed evidence suggests several cautious recommendations for policy and for the research–policy interface. Policymakers designing reform programs in LICs and LMICs should sequence structural, administrative, and distributive measures in ways that respect the tax-capacity ladder: building reliable registries, basic enforcement, and visible service improvements before attempting ambitious wealth or capital-gains taxation. Investments in digitalization are most likely to reinforce, rather than erode, the service–trust–revenue loop when they are paired with credible governance reforms, simplified procedures, and responsive taxpayer services. International partners and domestic researchers can enhance the external validity of policy advice by systematically documenting implementation details and context, and by making replication datasets—such as firm-level profitability and VAT-shock panels—available where legal and ethical frameworks permit. Addressing the empirical and methodological deficits identified in this review would help align normative ambitions of fairness, efficiency, and

legitimacy with implementable pathways to robust, context-appropriate revenue systems in low- and middle-income settings.

REFERENCES

Addison, T., Niño-Zarazúa, M., & Pirttilä, J. (2018). Fiscal Policy, State Building and Economic Development. *Journal of International Development*, *30*(2), 161–172. https://doi.org/10.1002/jid.3355

Agbanyo, R., Musah, G., Doku, J. N., Quarshie, J., & Dery, N. Y. (2024). Perceived political corruption and tax compliance intentions among private enterprises in emerging economies: The roles of tax morale and tax tolerance in Ghana. *Journal of International Development*, *36*(4), 2063–2095. https://doi.org/10.1002/jid.3894

Aizenman, J., Jinjarak, Y., Nguyen, H. T. K., & Park, D. (2019). Fiscal space and government-spending and tax-rate cyclicality patterns: A cross-country comparison, 1960–2016. *Journal of Macroeconomics*, *60*, 229–252. https://doi.org/10.1016/j.jmacro.2019.02.006

Akpoyibo, A. G., & Okolo, M. N. (2023). Tax Collection, Utilization Systems and the Performance of Small and Medium Scale Enterprises in Nigeria: A Multivariate Approach. *Journal of Tax Reform*, *9*(2), 134–144. https://doi.org/10.15826/jtr.2023.9.2.133

Ales, L., Bellofatto, A. A., & Wang, J. J. (2017). Taxing Atlas: Executive compensation, firm size, and their impact on optimal top income tax rates. *Review of Economic Dynamics*, *26*, 62–90. https://doi.org/10.1016/j.red.2017.02.007

Aliakbari Nouri, F., & Shafiei Nikabadi, M. (2025). Exploring the causal relationships between factors affecting taxpayer adoption of e-invoicing: application of interval neutrosophic DEMATEL. *Kybernetes*. https://doi.org/10.1108/K-07-2024-1807

Azémar, C., & Dharmapala, D. (2019). Tax sparing agreements, territorial tax reforms, and foreign direct investment. *Journal of Public Economics*, *169*, 89–108. https://doi.org/10.1016/j.jpubeco.2018.10.013

Bachas, P., Fattal Jaef, R. N., & Jensen, A. (2019). Size-dependent tax enforcement and compliance: Global evidence and aggregate implications. *Journal of Development Economics*, *140*, 203–222. https://doi.org/10.1016/j.jdeveco.2019.05.001

Berawi, M. A., Basten, V., Latief, Y., & Crévits, I. (2020). Role of green building developer and owner in sustainability construction: investigating the relationships between green building key success factors and incentives. *IOP Conference Series: Earth and Environmental Science*, *426*(1), 012061. https://doi.org/10.1088/1755-1315/426/1/012061

Bergolo, M., Londoño-Vélez, J., & Tortarolo, D. (2023). Tax progressivity and taxing the rich in developing countries: lessons from Latin America. *Oxford Review of Economic Policy*, *39*(3), 530–549. https://doi.org/10.1093/oxrep/grad029

Castañeda Rodríguez, V. M. (2018). Tax determinants revisited. An unbalanced data panel analysis. *Journal of Applied Economics*, *21*(1), 1–24. https://doi.org/10.1080/15140326.2018.1526867

Cevik, S., & Miryugin, F. (2022). Death and taxes: Does taxation matter for firm survival? *Economics & Politics*, *34*(1), 92–112. https://doi.org/10.1111/ecpo.12188

Davies, M. H., & Schröder, M. (2025). A simple model of internal and external balance for resource-rich developing countries. *Australian Journal of Agricultural and Resource Economics*, *69*(1), 178–199. https://doi.org/10.1111/1467-8489.12589

Doherty, J., Kirigia, D., Okoli, C., Chuma, J., Ezumah, N., Ichoku, H., Hanson, K., & McIntyre, D. (2018). Does expanding fiscal space lead to improved funding of the health sector in developing countries?: lessons from Kenya, Lagos State (Nigeria) and South Africa. *Global Health Action*, *11*(1), 1461338. https://doi.org/10.1080/16549716.2018.1461338

Elliott, L. M., Dalglish, S. L., & Topp, S. M. (2020). Health Taxes on Tobacco, Alcohol, Food and Drinks in Low- and Middle-Income Countries: A Scoping Review of Policy Content, Actors, Process and Context. *International Journal of Health Policy and Management*. https://doi.org/10.34172/ijhpm.2020.170

Gnangnon, S., & Brun, J. (2017). Impact of export upgrading on tax revenue in developing and high-income countries. *Oxford Development Studies*, *45*(3), 388–408. https://doi.org/10.1080/13600818.2016.1278748

Gwaindepi, A. (2021). Domestic revenue mobilisation in developing countries: An exploratory analysis of sub-Saharan Africa and Latin America. *Journal of International Development*, *33*(2), 396–421. https://doi.org/10.1002/jid.3528

Jackson, A., & Sanger, T. (2020). Policy Forum: The Case for an Annual Net Wealth Tax. *Canadian Tax Journal/Revue Fiscale Canadienne*, *68*(3), 835–850. https://doi.org/10.32721/ctj.2020.68.3.pf.jackson

Kabir, M. (2021). Behavioural intention to adopt blockchain for a transparent and effective tax management system. *Journal of Global Operations and Strategic Sourcing*, *14*(3), 522–544. https://doi.org/10.1108/jgoss-09-2019-0077

Kanbur, R., Paukkeri, T., Pirttilä, J., & Tuomala, M. (2018). Optimal taxation and public provision for poverty reduction. *International Tax and Public Finance*, *25*(1), 64–98. https://doi.org/10.1007/s10797-017-9443-6

Karakotsios, A., Katrakilidis, C., Dimitriadis, D., & Christoforidis, T. (2020). Examining the relationship between income inequality, taxation and economic freedom: a panel cointegration approach. *Economics and Business Letters*, *9*(3), 206–215. https://doi.org/10.17811/ebl.9.3.2020.206-215

Kebede, T. N., Erena, O. T., & Bawiso, E. P. (2024). Determinants of Tax Revenue: A Cointegration and Causality Analysis for Ethiopia, 1992–2022. *Journal of Tax Reform*, *10*(3), 493–509. https://doi.org/10.15826/jtr.2024.10.3.180

Kraus, S., Breier, M., Lim, W. M., Dabić, M., Kumar, S., Kanbach, D., Mukherjee, D., Corvello, V., Piñeiro-Chousa, J., Liguori, E., Palacios-Marqués, D., Schiavone, F., Ferraris, A., Fernandes, C., & Ferreira, J. J. (2022). Literature reviews as independent studies: guidelines for academic practice. *Review of Managerial Science*, *16*(8), 2577–2595. https://doi.org/10.1007/s11846-022-00588-8

Lim, W. M., Yap, S.-F., & Makkar, M. (2021). Home sharing in marketing and tourism at a tipping point: What do we know, how do we know, and where should we be heading? *Journal of Business Research*, *122*, 534–566. https://doi.org/10.1016/j.jbusres.2020.08.051

Lois, P., Drogalas, G., Karagiorgos, A., & Chlorou, A. (2019). Tax compliance during fiscal depression periods: the case of Greece. *EuroMed Journal of Business*, *14*(3), 274–291. https://doi.org/10.1108/EMJB-02-2019-0028

López, R. E., & Figueroa B., E. (2016). On the Nexus Between Fiscal Policies and Sustainable Development. *Sustainable Development*, *24*(4), 201–219. https://doi.org/10.1002/sd.1622

Mahmoud, N. B., Hamed, E. H., & Dwidar, M. S. (2025). Assessing the impact of e-government on control of corruption and government effectiveness: Developed and developing countries. *Asian Economic and Financial Review*, *15*(5), 740–755. https://doi.org/10.55493/5002.v15i5.5391

Marthalina, M., Dione, F., & Saribulan, N. (2025). Determining Factors of Digital Public Service Management Implementation Policy at Subdistrict Level in South Jakarta City. *Jurnal Manajemen Pelayanan Publik*, *9*(2), 251–275. https://doi.org/10.24198/jmpp.v9i2.59855

Martorano, B. (2018). Taxation and Inequality in Developing Countries: Lessons from the Recent Experience of Latin America. *Journal of International Development*, *30*(2), 256–273. https://doi.org/10.1002/jid.3350

Mawejje, J., & Francis Munyambonera, E. (2016). Tax Revenue Effects of Sectoral Growth and Public Expenditure in Uganda. *South African Journal of Economics*, *84*(4), 538–554. https://doi.org/10.1111/saje.12127

Mazur, H., Zinevych, O., Guseva, O., Babichev, A., & Valyvsky, S. (2025). The Impact of the Shadow Economy on the Stability of the Financial System of the State. *Salud, Ciencia y Tecnología - Serie de Conferencias*, *4*, 1423. https://doi.org/10.56294/sctconf20251423

Mgammal, M., & Al Matari, E. (2023). Developing a panel and time series database of value added tax rate increase impacts [Data set]. In *DataONE*. https://search.dataone.org/

Mohammed, H., & Tangl, A. (2023). Taxation Perspectives: Analyzing the Factors behind Viewing Taxes as Punishment—A Comprehensive Study of Taxes as Service or Strain. *Journal of Risk and Financial Management*, *17*(1), 5. https://doi.org/10.3390/jrfm17010005

Muguchu, J., Wawire, N. H., & Wambugu, A. (2021). Taxable capacity and effort of value-added tax in Kenya. *African Multidisciplinary Tax Journal*, *2021*(1), 189–210. https://doi.org/10.47348/AMTJ/2021/i1a11

Muhammad, I., Mohd Hasnu, N. N., Ibrahim, M. A., Abdul Hamid, S., & Mohd Hanefah, M. (2022). Trust in Government and Its Determinants: An Empirical Study of Public Acceptability for Carbon Tax in Malaysia. *Sustainability*, *14*(23), 15684. https://doi.org/10.3390/su142315684

Nkundabanyanga, S. K., Mvura, P., Nyamuyonjo, D., Opiso, J., & Nakabuye, Z. (2017). Tax compliance in a developing country. *Journal of Economic Studies*, *44*(6), 931–957. https://doi.org/10.1108/JES-03-2016-0061

Piancastelli, M., & Thirlwall, A. P. (2020). The Determinants of Tax Revenue and Tax Effort in Developed and Developing Countries: Theory and New Evidence 1996-2015. *Nova Economia*, *30*(3), 871–892. https://doi.org/10.1590/0103-6351/5788

Purnamasari, D., & Tahir, R. (2025). Efforts to Improve the Public Service Quality: Implementation of a Modern Electronic Tax System, Self-Efficacy, and Taxpayer Compliance. *Jurnal Manajemen Pelayanan Publik*, *9*(1), 141–167. https://doi.org/10.24198/jmpp.v9i1.58803

Ranjan, A. (2023). International Tax Law and Low- and Middle-Income Countries. In *The Oxford Handbook of International Tax Law* (pp. 123–142). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780192897688.013.8

Reeves, A., Gourtsoyannis, Y., Basu, S., McCoy, D., McKee, M., & Stuckler, D. (2015). Financing universal health coverage—effects of alternative tax structures on public health systems: cross-national modelling in 89 low-income and middle-income countries. *The Lancet*, *386*(9990), 274–280. https://doi.org/10.1016/S0140-6736(15)60574-8

Robinson, Z., & De Beer, J. (2021). Revisiting corporate income tax determinants in Southern Africa. *Development Southern Africa*, *38*(4), 564–579. https://doi.org/10.1080/0376835X.2020.1842175

Samusevych, Y., Rącka, I., Vasylieva, T., & Tóth, A. (2024). Environmental tax reform efficiency: Prerequisites and consequences. *Journal of International Studies*, *17*(4).

Schmoll, M. (2021). Weak Street-level Enforcement of Tax Laws: The Role of Tax Collectors' Persistent but Broken Public Service Expectations. *The Journal of Development Studies*, *57*(2), 209–225. https://doi.org/10.1080/00220388.2020.1779928

Schoeman, A., & de la Feria, R. (2019). Addressing VAT Fraud in Developing Countries: The Tax Policy-Administration Symbiosis. *Intertax*, *47*(Issue 11), 950–967. https://doi.org/10.54648/TAXI2019096

Sebele-Mpofu, F. Y., Mashiri, E., & Korera, P. (2025). Transfer Pricing Audit Challenges and Dispute Resolution Effectiveness in Developing Countries with Specific Focus on Zimbabwe. *Accounting, Economics, and Law: A Convivium*, *15*(2), 223–269. https://doi.org/10.1515/ael-2021-0026

Söderström, J., & Wangel, M. (2023). Public Channels of Access and State-Citizen Encounters in the Namibian Tax System. *Administration & Society*, *55*(2), 264–293. https://doi.org/10.1177/00953997221133503

Storonyanska, I., Patytska, K., Medynska, T., Benovska, L., Kliuchnyk, L., & Nestor, O. (2023). SYSTEMIC PREREQUISITES FOR SHAPING THE LOCAL BUDGETS' TAX REVENUE: THE CASE STUDY OF DEVELOPED COUNTRIES. *Financial and Credit Activity Problems of Theory and Practice*, *6*(53), 58–69. https://doi.org/10.55643/fcaptp.6.53.2023.4204

Thi Hai Do, H., Thi Hai Mac, Y., Thi Van Tran, H., & Thi Le Nguyen, T. (2022). The impact of attitude towards an e-tax system on tax compliance of Vietnamese enterprises: Adoption of an e-tax system as a mediator. *Journal of Entrepreneurship, Management and Innovation*, *18*(1), 35–64. https://doi.org/10.7341/20221812

Tikhonova, A., Telegina, Z., Babanskaya, A., & Grudneva, A. (2019). Evaluation of the optimal ratio for consumption and income taxes with the use of econometric methods. *Economic Annals-XXI*, *177*(5–6), 44–52. https://doi.org/10.21003/ea.V177-04

Tolkmitt, V., Volokhova, I., & Ignatenko, R. (2022). FINANCIAL DECENTRALIZATION IN UKRAINE IN THE CONTEXT OF EUROPEAN TRENDS. *Financial and Credit Activity Problems of Theory and Practice*, *4*(45), 36–47. https://doi.org/10.55643/fcaptp.4.45.2022.3839

Umar, M. A., Derashid, C., & Ibrahim, I. (2017). What Is Wrong With the Fiscal Social Contract of Taxation in Developing Countries? A Dialogue With Self-Employed Business Owners in Nigeria. *Sage Open*, *7*(4), 1–14. https://doi.org/10.1177/2158244017745114

Ventura-Cots, M., Ballester-Ferré, M. P., Ravi, S., & Bataller, R. (2019). Public health policies and alcohol-related liver disease. *JHEP Reports*, *1*(5), 403–413. https://doi.org/10.1016/j.jhepr.2019.07.009

von Haldenwang, C., & von Schiller, A. (2016). The Politics of Taxation: Introduction to the Special Section. *The Journal of Development Studies*, *52*(12), 1685–1688. https://doi.org/10.1080/00220388.2016.1153075

Wang, T., Wan, L., & Bao, H. X. H. (2025). Land-use efficiency and local government revenue: Evidence from 272 Chinese cities using a novel structural equation modelling approach. *Environment and Planning B: Urban Analytics and City Science*, *52*(4), 804–822.