NIPFP Working paper series

Bridging the Budget Divide: How PEFA Assessments and PFM Reforms Diagnose the Budget Reliability

No. 435
August-2025
Dinesh Kumar Nayak
Pratap Ranjan Jena

National Institute of Public Finance and Policy
New Delhi

Bridging the Budget Divide: How PEFA Assessments and PFM Reforms Diagnose the Budget Reliability

Dinesh Kumar Nayak and Pratap Ranjan Jena*

Abstract

In this paper, we examine how Public Financial Management (PFM) systems affect budget reliability, focusing on interconnected PEFA pillars, indicators, and dimensions, across the globe, mainly in lower-middle-income countries (LMICs). Using the PEFA Public Database 2022 Global Report, we identify that there is a widespread challenge in maintaining reliable budgeting globally, more pronounced in lower-income group countries, and many Indian states exhibit growing discrepancies between planned and actual revenues and expenditures. We employ structural equation modeling (SEM) to analyze the effects of PFM system on budget reliability, and measurement models to construct latent variables from the observed indicators, such as PEFA pillars and dimensions. We find that a substantial and significant positive standardized effect of each unit increase in PFM system rises budget reliability by 0.625 (globally) and 0.629 (LMICs) units. Based on effect-size estimates of parameters, external scrutiny and audits are vital for enhancing budget reliability in PFM systems globally and in LMICs. Additionally, transparency in public finances is crucial globally, while predictability and control in budget execution are essential for LMICs. The findings offer policy relevance for developing and emerging nations, including India.

Keywords: Public Financial Management (PFM); Public Expenditure and Financial Accountability (PEFA) Framework; Budget Reliability; Fiscal Discipline

JEL Classifications: M40; E62; H61; H70; H72; H83; O22

1 Introduction

An increasing number of policymakers, scholars, and practitioners recognize the critical link between public financial management (PFM) and governance and development outcomes (Cuadrado-Ballesteros et al., 2020). The integrity of PFM systems ensures the efficient budgeting, expenditure, and reporting of public resources. Among the 64 available diagnostic tools, the Public Expenditure and Financial Accountability (PEFA), launched in 2001 is by far the most complete and used instrument (Andrews et al. 2014, Kristensen et al. 2019). It allows the performance of the PFM to be measured through 94 dimensions in 31 performance indicators (PIs) grouped into seven key pillars (Table 1): P1 – Budget Reliability, P2 – Transparency in Public Finances, P3 – Management of Assets and Liabilities, P4 – Policy-Oriented Fiscal Strategy and Budgeting, P5 – Predictability and Control in Budget Execution, P6 – Accounting and Reporting, and P7 – External Scrutiny and Audit. These pillars provide a structured approach to managing public finances, improving service delivery, and achieving government objectives.

Budget evaluation and accountability are central to effective PFM, as the budget is a critical public document that expresses the priorities and commitments of a government, at the national or subnational level (UNDESA & IBP, 2023). It details planned revenue and expenditure to address competing national needs, such as security, healthcare, and poverty alleviation. Given its profound implications for citizens, the budget demands rigorous scrutiny and debate (Friedman, 2006). As

*Authors are respectively Economist and Professor at National Institute of Public Finance and Policy (NIPFP), New Delhi, India - 110067. An earlier version of this paper was presented at GIFT-MSE International Conference themed "Rethinking Public Finance for Emerging Development Challenges", held on 18-21 March 2025. We are grateful to the participants of the conference for their valuable observations, and comments. The usual disclaimer applies. Emails: dinesh.nayak@nipfp.org.in; pratap.jena@nipfp.org.in

countries intensify efforts to achieve the Sustainable Development Goals (SDGs), financing gaps highlight the urgency of robust PFM systems (de Renzio, Lakin & Cho, 2019). The PEFA framework (Figure 1) offers a systematic methodology for assessing budgeting systems, directly influencing fiscal discipline, resource allocation, and stakeholder trust.

Table 1 Composition of PEFA Pillars and Performance Indicators

Pillar(s)		Performance Indicator(s)	
1	Budget reliability	1	Aggregate expenditure outturn
		2	Expenditure composition outturn
		3	Revenue outturn
2	Transparency of public finances	4	Budget classification
		5	Budget documentation
		6	Central government operations outside financial reports
		7	Transfers to subnational gove
		8	Performance information for se
		9	Public access to fiscal information
3	Management of assets and liabilities	10	Fiscal risk reporting
		11	Public investment management
		12	Public asset management
		13	Debt management
4	Policy-based fiscal strategy and budgeting	14	Macroeconomic and fiscal forecasting
		15	Fiscal strategy
		16	Medium-term perspective in expenditure budgeting
		17	Budget preparation process
		18	Legislative scrutiny of budgets
5	Predictability and control in budget execution	19	Revenue administration
		20	Accounting for revenue
		21	Predictability of in-year resource allocation
		22	Expenditure arrears
		23	Payroll controls
		24	Procurement
		25	Internal controls on nonsalary expenditure
		26	Internal audit
6	Accounting and reporting	27	Financial data integrity
		28	In-year budget reports
		29	Annual financial reports
7	External scrutiny and audit	30	External audit
		31	Legislative scrutiny of audit reports

Source: Framework for assessing PFM, October 2019, PEFA Secretariat

The PEFA framework has played a pivotal role in assessing PFM systems, highlighting the connection between PFM reforms and broader economic development (de Renzio, 2009; Long, 2019). With approximately 800 assessments conducted in more than 155 countries, its global reach is substantial (Upadhaya et al., 2024). However, challenges persist in developing countries, where discrepancies between budgeted and actual figures often signal planning or execution issues (Grossi et al., 2020; Kristensen et al., 2019). In India, for instance, state governments exhibit poor fiscal accuracy in revenue and expenditure components (Jena & Singh, 2021; Chakraborty et al., 2020). Studies using the PEFA framework reveal deteriorating budget forecasts, with an increasing number of states scoring below basic levels and fewer achieving advanced scores (RBI, 2024). Only two PEFA assessments have been conducted in India, one national (Jena, 2010) and one sub-national in Himachal Pradesh (World Bank, 2009), highlighting a significant gap in PFM evaluation (PEFA Secretariat, 2022). This scarcity calls for dialogue among stakeholders to promote PEFA assessments, which can pinpoint strengths and weaknesses in PFM systems and guide reforms to improve budget reliability and fiscal accuracy.

1.1 Objective of the Study:

This study aims to investigate the role of the PEFA framework in strengthening budget evaluation and accountability within PFM systems, with a particular focus on lower-middle-income countries

(LMICs) and India. Specifically, it seeks to evaluate the total, direct, and indirect effects of a sound PFM system, as measured by the PEFA pillars, on budget accountability. Additionally, the study employs structural equation modeling (SEM) to empirically analyze the cause-and-effect relationships among the PEFA dimensions, performance indicators, and pillars, using data from the PEFA Public Database 2022 Global Report.

1.2 Research Questions:

The following research questions guide this investigation.

1. How does the PEFA framework improve budget reliability in PFM systems globally and in LMICs?
   This question explores the broader impact of the framework, drawing on its global application (Upadhaya et al., 2024) and its relevance to developing contexts (Kristensen et al., 2019).
2. What are the cause-and-effect relationships among the PEFA dimensions, performance indicators, and pillars, and how do they influence budget accountability?
   This question uses SEM to examine the internal dynamics of the framework, informed by studies on the interdependencies of the PFM system (Andrews et al., 2014).
3. What are the specific challenges and opportunities for implementing the PEFA framework in India to enhance budget reliability and fiscal precision?
   This focuses on India's unique PFM landscape, building on evidence of fiscal inaccuracies (Jena & Singh, 2021; RBI, 2024).

1.3 Hypotheses:

Drawing from the research questions and existing literature, the study tests the following hypotheses.

- **H1**: A sound PFM system, as indicated by higher scores in the PEFA pillars, positively influences budget accountability.
  This hypothesis is supported by evidence linking strong PFM systems to enhanced fiscal discipline and transparency (de Renzio, 2009; Cuadrado-Ballesteros et al., 2020).
- **H2**: The relationship between specific PEFA pillars and budget accountability is mediated by other pillars or indicators.
  The interconnected nature of the PEFA pillars suggests mediation effects, where improvements in transparency or control mechanisms can indirectly boost the budget reliability performance (Kristensen et al., 2019).

Fig. 1 Interrelationship of the PEFA pillars of the PFM system in Budget cycle

Source: Framework for assessing PFM, October 2019, PEFA Secretariat

To address these questions and hypotheses, the study employs SEM, a statistical technique suited for analysing complex relationships among observed and latent variables (Flora, Crone, & Bell, 2025; Kline, 2023). Using data from the PEFA Public Database 2022 Global Report, this study employs SEM to construct structural and measurement models for budget evaluation and accountability. Path diagrams are developed and interpreted to examine direct and indirect effects within the PEFA framework, contributing to both policy insights and academic literature on PFM. This study provides a robust foundation for exploring the PEFA framework's impact on budget evaluation and accountability, offering actionable insights for LMICs and India while advancing PFM scholarship through empirical analysis.

The remainder of this paper is organized as follows: Section 2 engages in a discussion of the extant literature and elucidates the existing research gap. Section 3 offers a comprehensive overview of the data and delves into the econometric methodology, outlining the identification strategy. Section 4 articulates the empirical findings. Finally, Section 5 concludes by summarizing the study and proposing some policy recommendations.

2 Literature Review: PFM, PEFA, and SEM

PFM systems are crucial for effective governance, fiscal discipline, and public accountability. However, there is a notable research gap in understanding how the strength of PFM systems directly impacts budget accountability, especially in aligning budgets with approved allocations. Budget accountability, characterized by minimal deviations from planned expenditures in total and composition (Mustapha, 2019), is fundamental to maintaining fiscal credibility and public trust. While existing studies offer some evidence of a link between robust PFM systems and enhanced budget credibility, the findings are inconsistent, and few have utilized advanced statistical methods, such as regression methods, to explore these dynamics rigorously. This study seeks to address this gap by employing SEM to evaluate the cause-and-effect relationships within the PEFA framework, focusing on its impact on budget accountability in low- and middle-income countries (LMICs).

2.1 Existing Evidence: PFM and Budget Accountability

The literature provides mixed but generally supportive evidence of the connection between PFM quality and budget accountability. Addison (2013) analyzed expenditure deviations using PEFA reports across 56 countries, identifying a correlation between accurate budget composition and higher PFM quality. However, the relationship between PFM quality and overall expenditure deviation was less pronounced, suggesting that additional variables may mediate this link. Similarly, Fritz, Sweet, and Verhoeven (2014) established a significant positive association between PFM quality and overall budget accountability, though their analysis did not account for potential confounding factors.

More recent research builds on these insights. Mustapha (2019), as cited in Robinson et al. (2021), examined 116 countries using PEFA assessments and found that strengthening PFM systems can enhance budget credibility, even in fragile states, by reducing variance in expenditure composition. This suggests that PFM reforms may stabilize budget execution across diverse contexts. However, Omollo's (2018) comparative case study of Kenya and Rwanda highlights the role of contextual factors: while PFM reforms improved revenue outcomes in Rwanda, similar efforts in Kenya yielded limited results, pointing to the influence of policy and institutional environments.

From a broader perspective, De Lay et al. (2015) conducted a systematic review of 197 studies on PFM interventions in LMICs. Their findings indicate that such interventions frequently enhance fiscal transparency (75 instances), public accountability (60 instances), and resource allocation alignment with policy objectives (52 instances). These outcomes highlight the governance benefits of PFM systems, although their effectiveness depends on the context. Similarly, Upadhaya et al. (2024) found that better PFM quality leads to increased fiscal transparency and public accountability. They also noted that institutional factors partially moderate these effects, indicating that PFM impacts are not consistently uniform.

2.2 Functional Perspective and Contextual Influences

From a functional perspective, the public accountability framework proposed by Bovens (2007, 2010) suggests that the quality and extent of PFM systems should enhance accountability and transparency. However, accounting and public sector literature emphasizes that these relationships are contingent on institutional contexts (Adhikari et al., 2023). Studies on specific PFM reforms, such as program-based

budgeting, integrated financial management information systems (IFMIS), and accrual accounting, often focus on technical implementation, but acknowledge that broader institutional settings shape their success. For instance, effective adoption of IFMIS may improve transparency in one country but falter in another due to weak governance structures. This context-dependency highlights the need for a nuanced approach to understanding PFM outcomes.

2.3 Role of SEM in PEFA and PFM Evaluation

Despite the growing body of research on PFM, advanced analytical techniques – particularly advanced regression methods – remain underutilized in examining systematic relationships. SEM analyzes latent variables and estimates effects, enhancing insights into how PFM components influence budget accountability. Coupling SEM with the PEFA framework, which includes dimensions, indicators, and pillars, is a robust approach to comprehending these interactions.

Despite extensive inquiries, no research has been identified that directly utilizes SEM with PEFA data for PFM evaluation. This dearth suggests that the application of SEM in this specific context is underdeveloped, possibly because of the complexity of PEFA data or the novelty associated with the combination of these methodologies. Existing research indicates a limited number of published studies directly employing SEM for PFM and PEFA evaluation, with an absence of studies explicitly combining these methodologies.

Nevertheless, the application of SEM in related domains, such as fiscal policy analysis (Thach, Oanh & Chuong 2018), underscores its potential in analyzing PEFA data. The available evidence suggests that SEM is a promising tool for future research, especially considering PEFA's comprehensive dataset and the pressing need for sophisticated statistical techniques to model intricate PFM relationships.

2.4 Advancing SEM: PEFA for Budget Accountability

This research examines the interactions, structure, and global implementation of PFM mechanisms, with a particular emphasis on LMICs. By employing SEM on PEFA data, the study aims to elucidate the causal relationships among PFM components and their aggregate effect on budget accountability. This method addresses the shortcomings of previous studies by offering a comprehensive and statistically rigorous analysis that considers mediating and moderating variables. The empirical results are anticipated to provide actionable insights for the optimization of PFM systems to enhance budgetary evaluation and accountability, especially in contexts with limited resources.

By integrating existing evidence, recognizing contextual factors, and introducing SEM as an innovative analytical approach, this study seeks to advance the comprehension of the role of PFM systems in strengthening fiscal governance. The emphasis on LMICs ensures applicability to nations where PFM reforms are essential, thereby facilitating more effective budget administration and public accountability.

3 Data and Estimation Approach

This study adopts a two-stage approach to investigate the PEFA framework's role in strengthening PFM systems, with a focus on budget accountability and reliability.

In the first stage, the study addresses Research Questions 1 and 2 using SEM to analyze complex relationships among the process-oriented observed variable (PEFA pillars and indicators), and latent variables (PFM, and budget reliability). The analysis draws on data from the PEFA Public Database 2022 Global Report, specifically utilizing 122 assessment reports that adhere to the updated PEFA 2016 methodology. These reports, covering both national (67) and subnational (55) evaluations, were conducted and published between 2015 and 2021.[1]

The dataset reflects a diverse sample across income groups:

- High-Income Countries (HIC): 7 reports (5.74% of the sample)
- Low-Income Countries (LIC): 27 reports (22.13%)
- Lower-Middle-Income Countries (LMIC): 49 reports (40.16%)
- Upper-Middle-Income Countries (UMIC): 39 reports (31.97%)

[1] Notably, 91.8% of these assessments achieved a PEFA check certificate rating, affirming their methodological rigor and reliability.

This distribution enables a comparative analysis of PFM performance across economic contexts, and this study focuses on global and LMIC. Through SEM, the study constructs structural and measurement models to examine the direct and indirect effects of the PEFA pillars on budget accountability, testing Hypotheses 1 and 2. Hypothesis 1 posits that higher PEFA pillar scores positively influence budget accountability, while Hypothesis 2 explores whether specific pillars mediate the effects of others, reflecting the interconnected nature of PFM components.

3.1 Contextual Analysis with a Focus on India

In the second stage, the study addresses Research Question 3 by leveraging significant findings from the SEM analysis to identify gaps and opportunities within India's PFM landscape. Building on the quantitative insights, this phase examines the applicability of the PEFA framework in enhancing budget reliability and fiscal accuracy in India, where only two assessments — one national (Jena, 2010) and one subnational - Himachal Pradesh (World Bank, 2009) have been conducted. This limited application highlights a critical need for further evaluation to address persistent challenges, such as discrepancies between budgeted and actual figures at the state level.

3.2 Hypotheses Testing:

The SEM analysis in Stage 1 tests the following hypotheses:

- **H1**: Stronger PFM systems, as reflected by higher PEFA pillar scores, positively impact budget accountability.
- **H2**: The relationship between individual PEFA pillars and budget accountability is mediated by other pillars or performance indicators.

These hypotheses are evaluated through path diagrams and statistical outputs, providing empirical evidence to inform both policy and practice.

3.3 SEM Analysis for PFM and Budget Reliability

Structural Equation Modeling is a multivariate method used to explain complex relationships between observed and latent variables. It is particularly useful for analyzing PFM systems with latent constructs like PFM and budget reliability. SEM combines factor analysis and regression to explore direct and indirect effects, enhancing understanding of PEFA dimensions and indicators.

To investigate budget reliability, this study employed SEM with baseline specifications that incorporate both observed and latent variables. The observed endogenous variable, P1 (Budget Reliability), represents the outcome of interest, while the observed exogenous variables include P2 (Transparency in Public Finances), P3 (Management of Assets and Liabilities), P4 (Policy-Oriented Fiscal Strategy and Budgeting), P5 (Predictability and Control in Budget Execution), P6 (Accounting and Reporting), and P7 (External Scrutiny and Audit). These variables collectively inform the analysis of PFM processes.

This study utilizes **Confirmatory Factor Analysis (CFA)** measurement models to define the latent constructs within the SEM framework. CFA tests whether observed variables reliably measure underlying latent variables (Kline, 2023). Specifically, the latent exogenous variable **"pfm"** is constructed from the observed exogenous variables P2 through P7, capturing the multidimensional nature of PFM performance. The latent endogenous variable **"Budget-reliability"** is measured using three observed indicators: PI-1 (Aggregate Expenditure Outturn), PI-2 (Expenditure Composition Outturn), and PI-3 (Revenue Outturn). This structure allows us to assess the impact of "pfm" on "Budget-reliability".

The SEMs underwent iterative refinement and were finalized in accordance with **goodness-of-fit (GOF) criteria**, incorporating widely recognized metrics, such as the Root Mean Square Error of Approximation (RMSEA), Comparative Fit Index (CFI), and Tucker-Lewis Index (TLI) (Hu & Bentler, 1999). Path diagrams were developed and finalized utilizing Stata's SEM Builder, chosen due to its robust estimation and visualization capabilities. Reference **Figure 7** provides the path diagram for global public financial management (PFM), elucidating the relationships among variables within a broad context, whereas **Figure 8** represents the model adapted for PFM in low- and middle-income countries, underscoring contextual variations. Reference **Figure 9** illustrates the measurement model applied to both global and LMIC analyses.

Model validation involved testing the SEMs using **modification indices (MI)** and **stability analysis**. All MI values were below 3.841—the chi-square critical value for one degree of freedom at $p = 0.05$—indicating that no significant improvements could be achieved by adding paths (Brown, 2015). Stability analysis confirmed that all eigenvalues lay within the unit circle, satisfying the stability condition for simultaneous equation systems.

For the measurement models, **factor loadings** and **standardized estimates** were calculated to evaluate the strength and significance of the relationships between latent variables and their observed indicators. Building on this, this study estimated the **causal effect** of "pfm" on "Budget-reliability". Furthermore, **mediation analysis** was performed within multivariate regression path diagrams to estimate direct, indirect, and total effects for both global and LMIC contexts, providing a comprehensive view of the relationships.

In SEM, the **Wald test** tests hypotheses about model parameters like path coefficients, factor loadings, or variances by evaluating if they are significantly different from zero. This helps determine the significance of relationships or effects and the necessity of including paths in the model. Unlike goodness-of-fit indices (RMSEA, CFI, TLI), which evaluate overall model fit, the Wald test focuses on specific parameters. The Wald test in SEM assesses the statistical significance of paths between latent variables, such as "pfm" and "budget-reliability". A small p-value indicates the relationship should be retained in the model.

Residual analysis played a critical role in assessing model fit. Residuals, defined as the differences between observed and model-predicted values, were computed for all observed variables (Kline, 2023). Small residuals indicated a good fit, while residual patterns helped identify potential areas for refinement, enhancing the accuracy and interoperability of the model.

The SEM analyses rested on three key assumptions: **a large sample size, multivariate normality,** and **correct model specification**. While multivariate normality can sometimes be relaxed—e.g., through robust standard errors or bootstrapping—this study employed multiple GOF metrics to ensure no critical variables were omitted (Kline, 2023). SEM relies on asymptotic properties, necessitating a sufficiently large sample size for reliable parameter estimates. Various guidelines suggest sample sizes ranging from a minimum of 100 to ratios of observations to free parameters between 5:1 and 20:1. To determine the adequacy of the sample, this study performed a **statistical power analysis**, achieving a power of at least 0.8 to detect the estimated effects, consistent with standard recommendations (Cohen, 1992; Bowen & Guo, 2011). This approach aligns with techniques used in tools like pwrSEM (Wang & Rhemtulla, 2021), confirming the robustness of our sample size. The SEM analyses offer valuable insights into how PFM processes influence budget reliability, with implications for strengthening accountability in both global and LMIC settings.

4 Empirical Findings

Consequently, this study investigates how various PFM mechanisms and measures interact, are structured, and are implemented globally, with a particular focus on lower- and middle-income countries. The empirical findings of this study aim to provide deeper insights into how these PFM factors contribute to overall budget evaluation and accountability.

4.1 Distribution of Budget Reliability Performance across Income Group

The pie charts (**Figure 2**) illustrate the distribution of the budget reliability performance across different income groups—High-Income Countries (HIC), Low-Income Countries (LIC), Low-and-Middle-Income Countries (LMIC), Upper-Middle-Income Countries (UMIC)—and provide a total aggregated distribution. Performance levels are categorized as follows: A (advanced), B (better), C (basic good practices), and D (below basic). Analyzing these distributions reveals a clear pattern tied to income levels, with notable variations across groups.

- **High-Income Countries (HIC)**: 42.86% achieve advanced (A) performance, and 57.14% reach better (B) performance, with no countries in basic (C) or below-basic (D) categories. HIC exhibit the strongest budget reliability, with all countries performing at advanced or better levels. This suggests that wealthier nations possess robust financial management systems, likely supported by greater resources, strong institutional capacity, and effective governance structures.
- **Low-Income Countries (LIC)**: 48.15% are below basic (D), 33.33% at basic good practices (C), 18.52% at better (B), and none at advanced (A). Nearly half of LIC fall into the below-basic category, indicating significant challenges in budget reliability. This reflects systemic issues such

as limited financial controls, weak institutional frameworks, and possibly political instability, with no countries reaching advanced performance.

- **Low-and-Middle-Income Countries (LMIC)**: 40.82% at basic good practices (C), 26.53% at below basic (D), 24.49% better (B), and 8.16% at advanced (A). LMIC show a mixed performance, with the largest share at basic good practices but a notable presence in both below-basic and better categories. This variability suggests that while some LMIC are progressing, others face persistent obstacles to achieving reliable budgeting.
- **Upper-Middle-Income Countries (UMIC)**: 56.41% below basic (D), 17.95% at basic good practices (C), 17.95% at better (B), and 7.69% at advanced (A). Surprisingly, over half of UMIC exhibit below-basic performance, despite their higher income levels. This may point to inefficiencies, corruption, or governance gaps that undermine budget reliability, with only a small fraction reaching advanced levels.
- **Total Aggregated Distribution**: 39.34% below basic (D), 29.51% at basic good practices (C), 22.95% at better (B), and 8.20% at advanced (A). Globally, a significant portion of countries (39.34%) fall into the below-basic category, while only 8.20% achieve advanced performance. This highlights a widespread challenge in maintaining reliable budgeting, particularly pronounced in lower-income groups.

The data reveals a gradient of budget reliability performance closely linked to income levels. HIC significantly demonstrate superior performance, with no instances of below-basic outcomes, while LIC and UMIC struggle with high proportions of below-basic performance. LMIC present a more balanced but still lower-performing distribution. This suggests a strong correlation between economic development and the ability to maintain reliable budgets, though UMIC's poor performance indicates that income alone does not guarantee fiscal discipline.

Similar disparities are observed for other pillars (see **Figures 3** and **4**) among the different income groups.

4.2 Sign, Size, and Association among PEFA pillars focusing process and outcome

Global PFM:

Budget reliability exhibits moderate positive associations with most PEFA pillars (**Figure 5**), with the strongest link to External scrutiny and audit (0.48), followed by Transparency of public finances (0.44) and Fiscal strategy and budgeting (0.42). The weakest association is with Management of assets and liabilities (0.15), indicating it has minimal impact. Scatter plots confirm these relationships with upward trends for significant correlations and dispersed patterns for weaker ones. The histogram indicates that Budget reliability scores are slight right skew, with a longer tail toward lower scores, implying that while moderate performance is common, consistent with the performance of other pillars, but with some variability. The correlation coefficients, bolstered by statistical significance, suggest that enhancing external scrutiny, transparency, and fiscal strategy can meaningfully improve budget reliability, offering actionable insights for policymakers within the PEFA framework.

Sign: All correlation coefficients are positive, indicating that Budget reliability tends to improve as the scores of the other pillars increase. There are no negative associations, suggesting a general synergy among the pillars.

Size: The correlations range from weak (0.15) to moderate (0.48):

- The **strongest association** is with External scrutiny and audit (0.48), suggesting that effective oversight and audit processes significantly enhance budget reliability.
- **Moderate associations** exist with Transparency of public finances (0.44), Fiscal strategy and budgeting (0.42), Control in budget execution (0.39), and Accounting and reporting (0.31), indicating that these pillars also play a notable role in supporting budget reliability.
- The **weakest association** is with Management of assets and liabilities (0.15), implying that this pillar has minimal influence on budget reliability.

Association: The statistically significant correlations include those with External scrutiny and audit, Transparency of public finances, Fiscal strategy and budgeting, Control in budget execution, and Accounting and reporting. This suggests that improvements in these areas are reliably linked to

better budget reliability, while the weak, non-significant correlation with Management of assets and liabilities indicates a less consistent relationship.

The ***scatter plots*** in the correlation matrix visualize the relationships between Budget reliability and each of the other pillars:

- **Upward Trends**: For pillars with moderate correlations (e.g., External scrutiny and audit, Transparency of public finances, Fiscal strategy and budgeting, Control in budget execution, and Accounting and reporting), the scatter plots show upward-sloping trends. This confirms that higher scores in these pillars correspond to higher Budget reliability scores.
- **Weak Relationship**: The scatter plot for Budget reliability vs. Management of assets and liabilities is more dispersed, with no clear trend, aligning with the weak correlation (0.15). This suggests little to no linear relationship between these two pillars.
- **Variability**: Some scatter plots, such as Budget reliability vs. External scrutiny and audit, show a wider spread of points, indicating variability in how external oversight impacts budget reliability across different contexts. However, the overall positive trend remains evident.

The scatter plots thus provide visual confirmation of the positive associations and highlight the varying strengths of these relationships. The correlation coefficients, supported by statistical significance, provide a robust measure of how Budget reliability aligns with the other pillars, with External scrutiny and audit emerging as the most influential.

The scatterplot matrix reveals significant interdependencies among PEFA variables, with Predictability and Control in Budget Execution and Transparency of Public Finances showing the strongest relationship (0.64). Transparency of public finances and policy-based fiscal strategy emerges as a central pillar, influencing multiple dimensions, while Management of Assets and Liabilities and External Scrutiny and Audit exhibit weaker ties. These insights suggest that reforms targeting transparency and strategic budgeting could yield widespread benefits, whereas asset management and external oversight may need tailored approaches. This analysis provides a roadmap for policymakers to strengthen PFM systems effectively in case of global PFM.

LMIC PFM:

The PEFA scatter plot matrix (**Figure 6**) reveals a significantly interconnected budgeting core (Budget Reliability, Policy-Based Fiscal Strategy and Budgeting, Predictability and Control in Budget Execution, and External scrutiny and audit) with correlations of 0.35–0.46, supported by strongly links to Transparency of Public Finances with Predictability and Control in Budget Execution, and Accounting and Reporting (0.73–0.78). To promote PEFA assessments, LMICs should target reforms in the budgeting core for systemic impact, enhance transparency as an enabler, address weaker areas with specific interventions, and use benchmarking and donor support to drive progress. These strategies can elevate PFM performance, fostering economic stability and development in LMICs.

4.3 Path Diagrams Explaining Accountability Equations

SEM is utilized to define and analyze causal relationships within PFM systems through path diagrams. Path Diagrams - 1 and 2 (**Figure 7** and **8**) illustrate the causal relationships between variables specific to global PFM systems and LMIC, respectively. These diagrams are designed to test Hypothesis 2, which likely explores the structural relationships among observed variables in these distinct contexts.

In contrast, Path Diagram - 3 (**Figure 9**) is applicable to both global and LMIC analyses. It incorporates two latent variables:

- "pfm" (Public Financial Management), measured by observed exogenous variables P2 to P7, which correspond to key PFM dimensions such as Transparency in Public Finances, Management of Assets and Liabilities, and other PEFA pillars.
- "budget-reliability", assessed through observed performance indicators PI-1 to PI-3, representing metrics like Aggregate Expenditure Outturn, Expenditure Composition Outturn, and Revenue Outturn.

Path Diagram 3 (**Figure 9**) is specifically used to examine Hypothesis 1, which likely investigates the influence of PFM on budget reliability. Meanwhile, Path Diagrams 1 and 2 (**Figure 7** and **8**) validate Hypothesis 2 by modeling context-specific relationships.

The development and finalization of all path diagrams were accomplished through the utilization of robust Goodness-of-Fit (GOF) indices, including RMSEA, CFI, and TLI. These metrics are instrumental in ensuring that the models accurately represent the underlying data and offer a dependable foundation for the testing of hypotheses.

4.4 Goodness-of-fit

Table 2 provides Goodness-of-Fit (GOF) indices that evaluate how well the SEMs for budget accountability fit the observed data. These models include both Structural Models and Equivalent Measurement Models, assessed across two contexts: Global and LMIC.

1. **Chi-Square/Degree of Freedom (χ^2/df)**: This index measures the model's fit relative to its complexity. Values between 1 and 2 indicate a good fit. The Global Structural (1.247) and Global Measurement (1.569) models fall within this range, suggesting good fit. The LMIC models (0.387 and 0.939) are below 1, which is not ideal but can be acceptable if other indices are strong, as a value below 1 may indicate over-fitting or a very close fit to the data.
2. **Prob > Chi-Square_ms (Model Significance)**: A p-value greater than 0.05 indicates that the model's implied covariance matrix is not significantly different from the observed data, suggesting a good fit. Both the Global (0.273 and 0.086) and LMIC (0.944 and 0.511) for structural and measurement models exceed this threshold, strongly supporting their fit to the data.
3. **Prob > Chi-Square_bs (Baseline Significance)**: This index compares the model to a baseline (null or independence model). p-value of 0.000 indicates a poor fit of the baseline model, which is expected and appropriate in SEM. It suggests that there are significant relationships among the variables, supporting the application of a more detailed SEM to capture these relationships effectively.
4. **Root Mean Square Error of Approximation (RMSEA)**: RMSEA measures the discrepancy per degree of freedom. Values below 0.05 indicate a very good fit, and up to 0.08 is adequate. The Global Structural Model (0.045) and LMIC models (0.000) show very good to excellent fit, while the Global Measurement Model (0.069) is adequate. The LMIC models' RMSEA of 0.000 suggests an exceptionally close match to the data.
5. **Standardized Root Mean Square Residual (SRMR)**: SRMR represents the average residual difference between observed and predicted correlations. All values are well below 0.08, indicating a good fit across all models, with the LMIC Structural Model (0.024) showing the smallest residuals.
6. **Comparative Fit Index (CFI)**: CFI compares the model to a baseline model, with values closer to 1 indicating better fit. All models exceed 0.95, with the LMIC models achieving a perfect 1.000, demonstrating excellent fit. The Global models (0.995 and 0.983) also indicate strong support from the data.
7. **Tucker-Lewis Index (TLI)**: TLI adjusts CFI for model complexity. Values above 0.95 indicate good fit, though values exceeding 1 (as in LMIC models) can suggest overfit or an exceptionally good fit. All models meet or exceed 0.95, supporting their adequacy, with LMIC models showing particularly strong performance.
8. **Akaike's Information Criterion (AIC)**: AIC balances fit and complexity. Lower values suggest a better model. LMIC models have significantly smaller AIC values than their Global counterparts, indicating better fit or less complexity relative to the data.
9. **Bayesian Information Criterion (BIC)**: Like AIC, BIC penalizes complexity, with smaller values preferred. The LMIC models again outperform the Global models, reinforcing their superior fit.
10. **Overall Squared Multiple Correlation (SMC) / Equation-Level GFI**: SMC reflects the proportion of variance explained by the model. The Measurement Models (Global: 0.935, LMIC: 0.961) show excellent fit, indicating strong relationships between latent variables and observed indicators. The Structural Models have lower SMC (Global: 0.467, LMIC: 0.708), with the Global model explaining less variance, though the LMIC model performs well.
11. **Stability Index**: All eigenvalues inside the unit circle satisfies stability. A value of 0 indicates model stability, meaning no explosive behavior. Both structural and measurement models satisfy this condition, supporting their reliability.
12. **Wald Tests for Equations**: This test confirms that exogenous variables significantly influence endogenous variables in all models. The p-value of 0.000 across all equations strongly supports the hypothesized relationships in the SEMs.

Overall Assessment of Model Fit

- **Structural Models**:
  - **Global**: Most indices ($(\chi^2/df) = 1.247$, RMSEA = 0.045, SRMR = 0.039, CFI = 0.995, TLI = 0.984) indicate a good fit, though the moderate SMC (0.467) suggests less explained variance in endogenous variables.
  - **LMIC**: Excellent fit across all indices ($(\chi^2/df) = 0.387$, RMSEA = 0.000, SRMR = 0.024, CFI = 1.000, TLI = 1.126), with a higher SMC (0.708), strongly supporting the model.
- **Measurement Models**:
  - Global: Good fit ($(\chi^2/df) = 1.569$, RMSEA = 0.069, SRMR = 0.049, CFI = 0.983, TLI = 0.952), with a high SMC (0.935) indicating strong measurement of latent variables.
  - LMIC: Exceptional fit ($(\chi^2/df) = 0.939$, RMSEA = 0.000, SRMR = 0.036, CFI = 1.000, TLI = 1.101), with an excellent SMC (0.961).

The GOF indices collectively demonstrate that the SEMs for budget accountability are well-supported by the data in both Global and LMIC contexts. The **LMIC models** (both structural and measurement) exhibit particularly strong fit, with many indices reaching optimal levels (e.g., RMSEA = 0.000, CFI = 1.000). The **Global models** also show good fit, though the structural model's lower SMC suggests it explains less variance. Overall, the assumed relationships in these SEMs are robust, making them reliable tools for analyzing budget accountability across different economic contexts.

Further the **Modified Indices** for respective SEMs supports the model validation. Most MI values were below 3.841—the chi-square critical value for one degree of freedom at p = 0.05—indicating that no significant improvements could be achieved by adding paths. Moreover, the **Residual analysis** defined as the differences between observed and model-predicted values, were computed for all observed variables. The test results were smaller and close to zero that indicates a good fit, suggesting no further requirement in the enhancement of the models' accuracy.

4.5 Evaluation of the Measurement Model

Loadings of measurement model for global pfm

Table 3 provides factor loadings and standardized estimates for two latent variables: PFM system and Budget reliability. These values show how strongly each observed variable measures its respective latent construct. All standardized loadings are statistically significant indicating reliable measurement.

PFM System: shows that P2 (Transparency of public finances) Fixed at 1.000 (standardized: 0.777), serves as the reference indicator to set the scale. P5 (Predictability & control in budget execution) Highest standardized loading (0.811), explaining 65.7% of its variance (SMC = 0.657), making it the strongest indicator of the PFM system. P6 (Accounting & reporting) Strong loading (0.740), with 54.7% variance explained (SMC = 0.547). P7 (External scrutiny & audit) Lowest loading (0.414), with only 17.1% variance explained (SMC = 0.171), indicating a weaker but still significant relationship. Other indicators (P3: 0.676, P4: 0.659) show moderate to strong relationships, with SMC values of 0.457 and 0.435, respectively.

Budget Reliability: The observed variables are PI-1: Aggregate expenditure outturn, PI-2: Expenditure composition outturn, and PI-3: Revenue outturn. PI-1: Fixed at 1.000 (standardized: 0.758), serves as the reference indicator. PI-3: Strongest loading (0.741), though its SMC (0.390) suggests 39.0% variance explained. PI-2: Loading of 0.719, with 54.9% variance explained (SMC = 0.549), indicating a solid indicator. All three variables are strong, reliable measures of Budget reliability.

Covariances of Measurement Errors

The **table 3** also lists covariances (and standardized estimates) between measurement errors of certain observed variables, indicating shared sources of error or unmodeled factors. Negative covariances (e.g., P4 and P6, P3 and Budget reliability) suggest that when measurement error increases in one variable, it decreases in the other, possibly due to unmodeled factors. The positive covariance between P7 and Budget reliability indicates that their errors are influenced by similar unaccounted factors. These findings suggest potential model refinements, such as allowing correlated errors or adding

latent variables. However, the other GOF measures support overall good-fit gives a reason to go with the specified model for analysis.

4.6 Estimates of Structural Models

Table 4 provides the regression weights from the structural model, examining the relationship between the latent variables "PFM" (Public Financial Management) and "Budget reliability" in two contexts: Global and LMIC PFM.

In case of Global PFM, for each unit increase in PFM, Budget reliability increases by 1.181 units. The standardized estimate of 0.625 shows a strong positive relationship, with PFM explaining 39.0% of the variance in Budget reliability (SMC = 0.390). The result is highly significant ($p < 0.001$), confirming a robust link between PFM and Budget reliability globally.

For LMIC PFM, each unit increase in PFM, Budget reliability rises by 0.612 units. The standardized estimate of 0.629 indicates a strong positive relationship, nearly identical to the global model, with PFM explaining 39.6% of the variance in Budget reliability (SMC = 0.396). The relationship is significant ($p = 0.004$), though slightly less so than in the global context, possibly due to data variability.

Both models show a strong, significant positive effect of PFM on Budget reliability (standardized estimates: 0.625 vs. 0.629), with similar explanatory power (39.0% vs. 39.6%). However, the unstandardized effect is smaller in LMIC (0.612 vs. 1.181), suggesting a less pronounced absolute impact, potentially due to contextual differences.

Table 5 shows the structural model estimates for the Global PFM context, based on the provided regression weights and covariances.

a. **Regression Weights**: The model assesses how various PFM pillars (exploratory variables) influence endogenous variables, such as Budget Reliability (P1) and other pillars (P2–P7). The observed variables for Budget Reliability (P1) shows positive effect by Transparency of Public Finances (P2) (standardized estimate = 0.285, $p = 0.001$), indicating a significant contribution. The Policy-Based Fiscal Strategy & Budgeting (P4) is weak positive effect (0.145, $p = 0.098$), only marginally significant. While the External Scrutiny & Audit (P7) has strong positive effect (0.360, $p = 0.000$), with highly significant., the SMC: 0.366 (36.6% of variance explained), suggesting moderate explanatory power.

b. **Covariances of Exogenous Variables**: Significant positive correlations exist among the exogenous pillars. These covariances suggest interdependencies, particularly between Transparency (P2) and Fiscal Strategy (P4), and between Fiscal Strategy (P4) and External Scrutiny (P7).

Table 6 provides the regression weights from the structural model LMIC PFM, focusing on how exploratory variables influence the endogenous variables: Budget Reliability (P1), Predictability & Control in Budget Execution (P5), and External Scrutiny & Audit (P7).

In case of Budget Reliability (P1), the Predictability & Control in Budget Execution (P5) has a significant positive effect (standardized estimate = 0.351, $p = 0.004$), suggesting that stronger budget execution control enhances budget reliability. The External Scrutiny & Audit (P7) also shows a significant positive effect (standardized estimate = 0.360, $p = 0.003$), indicating that effective external oversight improves budget reliability. The SMC shows 33.4% of the variance in P1 is explained by these factors, reflecting moderate explanatory power.

For Predictability & Control in Budget Execution (P5), the Transparency of Public Finances (P2) exhibits a strong positive effect (standardized estimate = 0.534, $p = 0.000$), implying transparency is a major driver of budget execution control. The Management of Assets & Liabilities (P3) shows a moderate positive effect (standardized estimate = 0.339, $p = 0.005$), indicating asset and liability management supports execution predictability. SMC of 0.661, meaning 66.1% of the variance in P5 is explained, demonstrating high explanatory power.

For External Scrutiny & Audit (P7), the Policy-Based Fiscal Strategy & Budgeting (P4) has a weak positive effect (standardized estimate = 0.1269, $p = 0.047$), with marginal significance, suggesting a limited role in enhancing scrutiny. Accounting & Reporting (P6) displays a moderate positive effect (standardized estimate = 0.304, $p = 0.025$), indicating reliable accounting strengthens audit processes. SMC shows that 22.6% of the variance in P7 is explained, reflecting lower explanatory power.

4.7 Mediating effects of Structural Models

Table 7 provides the mediating effects of the SEM for global PFM. While the Budget Reliability has direct effect from the Transparency of Public Finances (0.377), Policy-Based Fiscal Strategy & Budgeting (0.204) and External Scrutiny & Audit (0.418) there is indirect effect enhanced by full mediation observed from Predictability & Control in Budget (0.381) and Accounting & Reporting (0.478). Total Effects is strongest for P6 on P2 (0.631) and P5 on P3 (0.639), showing interconnected PFM pillars with mediation enhancing impacts. Similarly, in case of LMIC PFM, Predictability & control in budget execution, and External scrutiny & audit has direct effects on Budget reliability respectively at 0.423 and 0.364 at the significance level of 1% error (**Table 8**). Indirect effects on budget reliability is seen from the Transparency of public finances, Accounting & reporting, Management of assets & liabilities, and Policy-based fiscal strategy. The strongest effect is from public transparency on predictability & control in budget execution in LMIC setting.

4.8 Understanding the Power Analysis

Table 9 provides the estimates of power analysis for the Structural Equation Models (SEMs) and confirm the robustness of the sample size for the estimated coefficients. The power values for both the structural and measurement models in the global (N=122) and LMIC (N=49) contexts, assesses the probability of detecting a true effect if it exists, with values ranging from 0 to 1. A power of 0.8 or higher is typically considered adequate, indicating an 80% or greater chance of correctly rejecting the null hypothesis, while values below 0.8 suggest a higher risk of missing a true effect (Type II error).

For the global structural model, most parameters have power values above 0.8, suggesting that N=122 is generally sufficient. However, exceptions like P1 $\sim$ P4 (0.55 and 0.34) and P7 $\sim$ P5 (standardized = 0.56) indicate that the sample size may not be robust for these specific relationships, particularly for standardized coefficients. The LMIC model, with N=49, shows more variability in power. Parameters like P1 $\sim$ P5 and P5 $\sim$ P2 have robust power, but others, such as P5 $\sim$ P3 and P7 $\sim$ P4 (especially standardized at 0.18), have low power, suggesting that the smaller sample size limits the ability to detect these effects reliably.

For the global measurement model, it assesses how well latent variables (e.g., P1, pfm) are measured by observed indicators (e.g., pi1, P2). Power values estimated for standardized coefficients are consistently 1 or very close (0.97), indicating that N=122 is fully robust for estimating these measurement relationships. In LMIC case, most power values are high (0.89–1), suggesting robustness. However, pfm =$\sim$ P7 (0.69) and P1 $\sim$ pfm (0.68) are below 0.8, indicating some weakness. It is noted that at a 10% alpha level (instead of 5%), these increase to 0.79 and 0.84, respectively, improving their detectability with a less stringent threshold.

5 Concluding Remarks

This study addresses the current gap in research concerning the relationship between the strength of PFM systems and budget accountability.

The distribution of budget reliability performance across income groups reveals a clear pattern: higher-income countries (HIC) excel with advanced and better outcomes, while lower-income groups (LIC, LMIC, UMIC) face significant challenges, with a global tendency toward below-basic performance. This gradient reflects broader economic disparities and has profound implications for stability, resource allocation, investor confidence, and development. Addressing these challenges requires targeted policy interventions—such as capacity building, governance reforms, and international support—to enhance fiscal management and promote equitable economic progress across all income levels.

The PFM system is best measured by P5 (Predictability & control) and P6 (Accounting & reporting), while Budget reliability is well-represented by PI-1, PI-2, and PI-3, all with strong, significant loadings. Significant error correlations, especially involving P3, P4, P6, P7, and Budget reliability, highlight areas where the model could be improved to account for shared influences.

In case of global PFM, the Budget Reliability (P1) is most strongly driven by External Scrutiny & Audit (P7) and Transparency (P2), with a weaker influence from Fiscal Strategy (P4). Predictability & Control (P5) is a critical pillar, significantly influencing P2, P3, P4, and P7, and is itself strongly shaped by Accounting & Reporting (P6). Accounting & Reporting (P6) consistently impacts multiple pillars (P2, P5, P7). The model explains moderate to high variance for most pillars (e.g., 45.0% for

P2), but less for P7 (13.8%). The exogenous pillars (P2, P3, P4, P7) are interrelated, reflecting a connected PFM system.

In case of LMIC, Budget Reliability (P1) is significantly shaped by both Predictability & Control (P5) and External Scrutiny & Audit (P7). Predictability & Control (P5) is strongly influenced by Transparency (P2) and moderately by Asset Management (P3). External Scrutiny & Audit (P7) is weakly to moderately affected by Fiscal Strategy (P4) and Accounting & Reporting (P6). The model explains variance best for P5 (66.1%), followed by P1 (33.4%), and least for P7 (22.6%).

5.1 Limitations:

The sample size is robust for most structural and all measurement model parameters (global context), with power values often at or near 1. Exceptions in the structural model (e.g., P1 $\sim$ P4, P7 $\sim$ P5 standardized) have lower power, suggesting that a larger sample might be needed for these specific coefficients. In LMIC context, the smaller sample size is sufficient for some structural parameters (e.g., P1 $\sim$ P5, P5 $\sim$ P2) and most measurement parameters, but others (e.g., P5 $\sim$ P3, P7 $\sim$ P4) have low power, particularly for standardized coefficients. This indicates reduced robustness compared to the global model. Both sample sizes strongly support the measurement models, while the structural models reveal limitations, especially in LMIC. Increasing the sample size or adjusting the alpha level (e.g., to 10%) could enhance power for weaker parameters, improving the reliability of the estimated coefficients.

PEFA scores performance indicators using multiple dimensions, each rated A to D based on specific criteria. All requirements must be fully met for a score; partial fulfilment lowers the score. C indicates basic performance as per global standards, while D shows sub-basic performance or lack of data. Since PEFA data are available in letter scores, this paper followed de Renzio (2009) and Mustapha (2019) among others, in converting the grading into numerical values (e.g., A = 4, B+ = 3.5, B = 3, C+ = 2.5, C = 2, D+ = 1.5, and D = 1), with higher values denoting better quality PFM of a country. Averaging PEFA performance indicator scores to construct pillars and assessing PFM across countries is limited by the assumption that all indicators carry equal weight despite varying importance and practical challenges such as inconsistent assessment timing.

5.2 Policy Discussions:

The stark contrast between HIC and lower-income groups (LIC, LMIC, UMIC) underscores global economic inequality. Without interventions, LIC's persistent challenges and UMIC's unexpected struggles could widen this gap, perpetuating cycles of underdevelopment. HIC's superior fiscal management reinforces their developmental advantage, widening disparities unless lower-income countries improve.

The data highlights the need for tailored interventions: LIC require urgent support to build institutional capacity, strengthen financial controls, and address systemic issues. UMIC need reforms to tackle inefficiencies and governance gaps despite higher income levels. LMIC could benefit from targeted assistance to move beyond basic practices toward better or advanced performance. HIC can serve as models, potentially sharing best practices through international cooperation.

Key lessons for India:

As a lower-middle-income country, India faces challenges in budget reliability and PFM systems, similar to others in its income group. The study shows higher-income countries outperform lower-income ones in budget reliability, highlighting the need for reforms. India can strengthen its PFM systems and improve budget accountability by focusing on predictability, control, transparency, and scrutiny while enhancing asset management and fiscal strategies. A holistic approach, international collaboration, and capacity building are essential for equitable economic progress and aligning with higher-income countries, ultimately boosting investor confidence and supporting development.

References

[1] Addison, D. (2013). The Quality of Budget Execution and Its Correlates. Policy Research *Working Paper* 6657, World Bank, Washington, DC.

[2] Adhikari, P., Nkundabanyanga, S., Soobaroyen, T. & Jayasinghe, K., (2023). Public Sector Accounting in Emerging Economies in the Evolving Post-Covid Era, *Journal of Public Budgeting Accounting and Financial Management*, 35(3), 297-308.

[3] Andrews, M., Cangiano, M., Cole, N., de Renzio, P., Krause, P., & Seligmann, R. (2014). This is PFM. *CID Working Paper* No. 285, July 2014.

[4] Bovens, M. (2007). Analysing and assessing accountability: a conceptual framework. *European Law Journal*, 13(4), 447-468.

[5] Bovens, M. (2010). Two Concepts of Accountability: Accountability as a Virtue and as a Mechanism. *West European Politics*, 33(5), 946-967

[6] Bowen, N. K., & Guo, S. (2011). *Structural equation modeling*. Oxford University Press.

[7] Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (2nd ed.). The Guilford Press.

[8] Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing Structural Equation Models* (pp. 136-162). Sage Publications.

[9] Chakraborty, L., Chakraborty, P., and Shresha, R. (2020). Budget credibility of subnational governments: Analyzing the fiscal forecasting errors of 28 States in India. *Working Paper* No. 964, Levy Economics Institute of Bard College, Annandale-on-Hudson.

[10] Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112(1), 155–159.

[11] Cuadrado-Ballesteros, B., Citro, F. & Bisogno, M. (2020). The role of public-sector accounting in controlling corruption: An assessment of organisation for economic co-operation and development countries. *International Review of Administrative Sciences*, 86(4), 729-748.

[12] De Lay, S., Mills, L., Jadeja, K. & Lucas, B. (2015). *Public Financial Management Evidence Mapping*. Birmingham, UK: GSDRC, University of Birmingham.

[13] De Renzio, P. (2009). Taking Stock: What Do PEFA Assessments Tell Us about PFM Systems across Countries? *Working Paper 302*, Overseas Development Institute, London.

[14] De Renzio, P., Lakin, J., & Cho, C. (2019). *Budget credibility across countries: how deviations are affecting spending on social priorities*. Paper. Washington, DC: International Budget Partnership.

[15] Flora, D. B., Crone, G., & Bell, S. M. (2025). Effect Size Interpretation in Structural Equation Models. *Structural Equation Modeling: A Multidisciplinary Journal*, 13th Feb. 2025, 1-8.

[16] Fritz, V., Sweet, S. and Verhoeven, M. (2014). *Strengthening Public Financial Management: Exploring Drivers and Effects*. Washington, DC: World Bank.

[17] Grossi, G., Ho, A. T. & Joyce, P. G. (2020). Budgetary responses to a global pandemic: international experiences and lessons for a sustainable future. *Journal of Public Budgeting, Accounting & Financial Management*, 32(5), 737-744.

[18] Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6(1), 1–55.

[19] Jena, P. R. (2006). Fiscal Marksmanship: Link between Forecasting Central Tax Revenues and State Fiscal Management. *Economic and Political Weekly*, Volume 41, No. 37, September 16-22.

[20] Jena, P. R. (2010). India Public Expenditure and Financial Accountability: *Public Financial Management Performance Assessment Report* (March 2010).

[21] Jena, P. R. (2013). Improving public financial management in India: Opportunities to move forward. *NIPFP Working Paper Series*, No. 2013-123, New Delhi.

[22] Jena, P. R., and Sikdar, S. (2019). Budget Credibility in India: Assessment through PEFA Framework. *NIPFP Working Paper Series*, No. 284, New Delhi.

[23] Jena, P. R., and Singh, A. (2021). Sub-national Budget Credibility: Institutional Perspective and Reform Agenda in India. *NIPFP Working Paper Series*, No. 338, New Delhi.

[24] Friedman, J. (2006). *A Guide to Tax Work for NGOs*. International Budget Partnership (IBP).

[25] Kang, H., & Ahn, J. W. (2021). Model setting and interpretation of results in research using structural equation modeling: A checklist with guiding questions for reporting. *Asian nursing research*, 15(3), 157-162.

[26] Kline, R.B. (2023). *Principles and practice of structural equation modeling* (5th ed). New York: The Guilford Press.

[27] Kristensen, J. Bowen, M., Long, C., Mustapha, S. & Zrinski, U. (2019). *PEFA: Public Financial Management, and Good Governance*. The World Bank.

[28] Long, C. (2019). PFM and Perceptions of Corruption, In Kristensen, J. K., Bowen, M., Long, C., Mustapha, S. and Zrinski, U. *PEFA, Public Financial Management, and Good Governance*. World Bank.

[29] Mustapha, S. (2019). Budget Credibility, Fiscal Outcomes, and PFM Performance in Fragile and Nonfragile Countries, In Kristensen, J. K., Bowen, M., Long, C., Mustapha, S. and Zrinski, U. *PEFA, Public Financial Management, and Good Governance*. World Bank.

[30] Omollo, B. (2018). Public Financial Management Reforms and Fiscal Policy: A Qualitative Content Analysis of the Impact of PFM Reforms on Country Budget Income in Sub Saharan Africa: A Case for Kenya and Rwanda. *OECD Journal on Budgeting* 18 (1): 167–201.

[31] Nayudu, A. S.H. and Chakraborty, L. (2025). Fiscal Transparency and Budgetary Processes in India: Issues before the 16th Finance Commission. *NIPFP Working Paper*, No. 424, New Delhi.

[32] Reserve Bank of India. (2024). *State Finances: A Study of Budgets of 2024-25*, Fiscal Reforms by States. Reserve Bank of India.

[33] Robinson, T., McGrann, J., Kadir, A., Jiménez-Ontiveros, I., & Barton, U. A. (2021). Analyzing the Link between PFM System Strength and Public Investment Performance. *PEFA Research Paper*. Washington, DC: PEFA.

[34] Thach, N. N., Oanh, T. T. K., & Chuong, H. N. (2018). The Interaction Between Fiscal Policy, Macroprudential Policy and Financial Stability in Vietnam-An Application of Structural Equation Modeling. *Studies in Computational Intelligence*, 275–288.

[35] United Nations Department of Economic and Social Affairs (UNDESA) and the International Budget Partnership (2023). *Strengthening Budget Credibility through External Audits: A Handbook for Auditors*.

[36] Upadhaya, B., Soobaroyen, T., Ntim, C., Adhikari, P., Jayasinghe, K., Zalata, A., & Haque, F. (2024). To what extent do public financial management systems support fiscal transparency and public accountability? *PEFA Research Paper*. Washington, DC: PEFA.

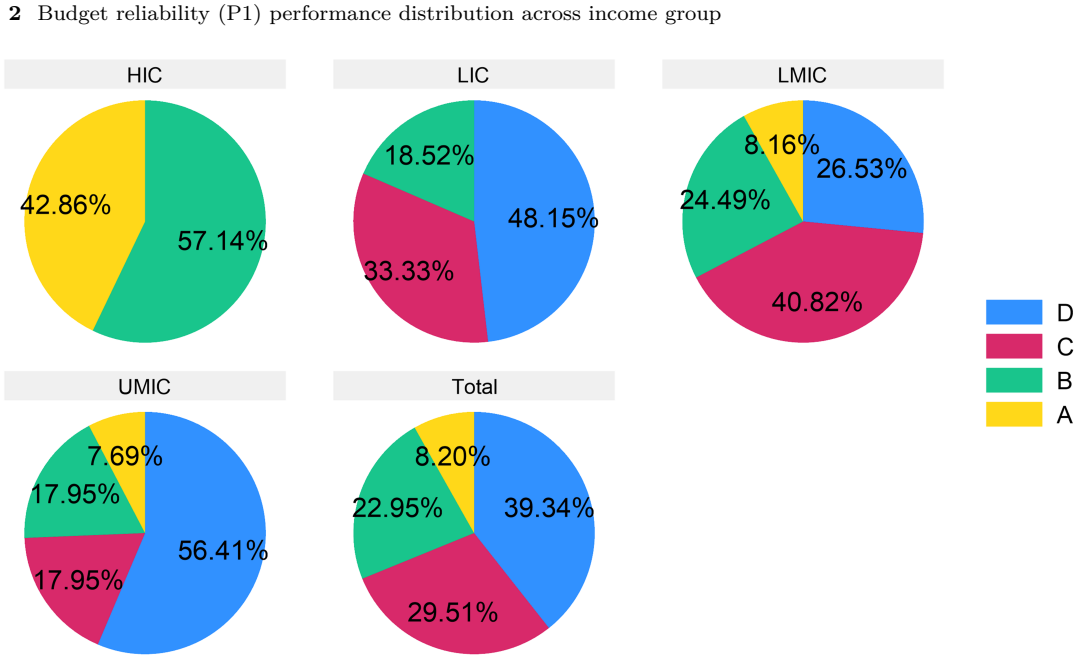
[37] Wang, Y. A., & Rhemtulla, M. (2021). Power analysis for parameter estimation in structural equation modeling: A discussion and tutorial. *Advances in Methods and Practices in Psychological Science*, 4(1), 2515245920918253.

[38] Website: PEFA: https://www.pefa.org

[39] World Bank (2009). India–Himachal Pradesh Public Financial Management Accountability Assessment *Public Expenditure and Financial Accountability Assessment Report* (Report No. 48635-IN, June 2009, Financial Management Unit, South Asia Region).

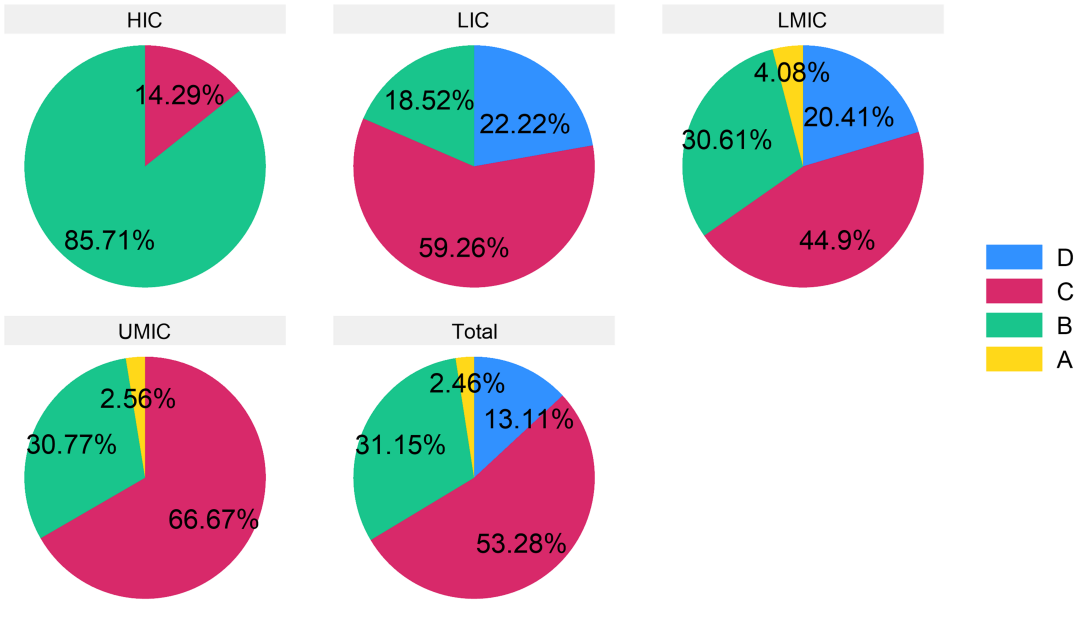
Figure(s):

Fig. 2 Budget reliability (P1) performance distribution across income group



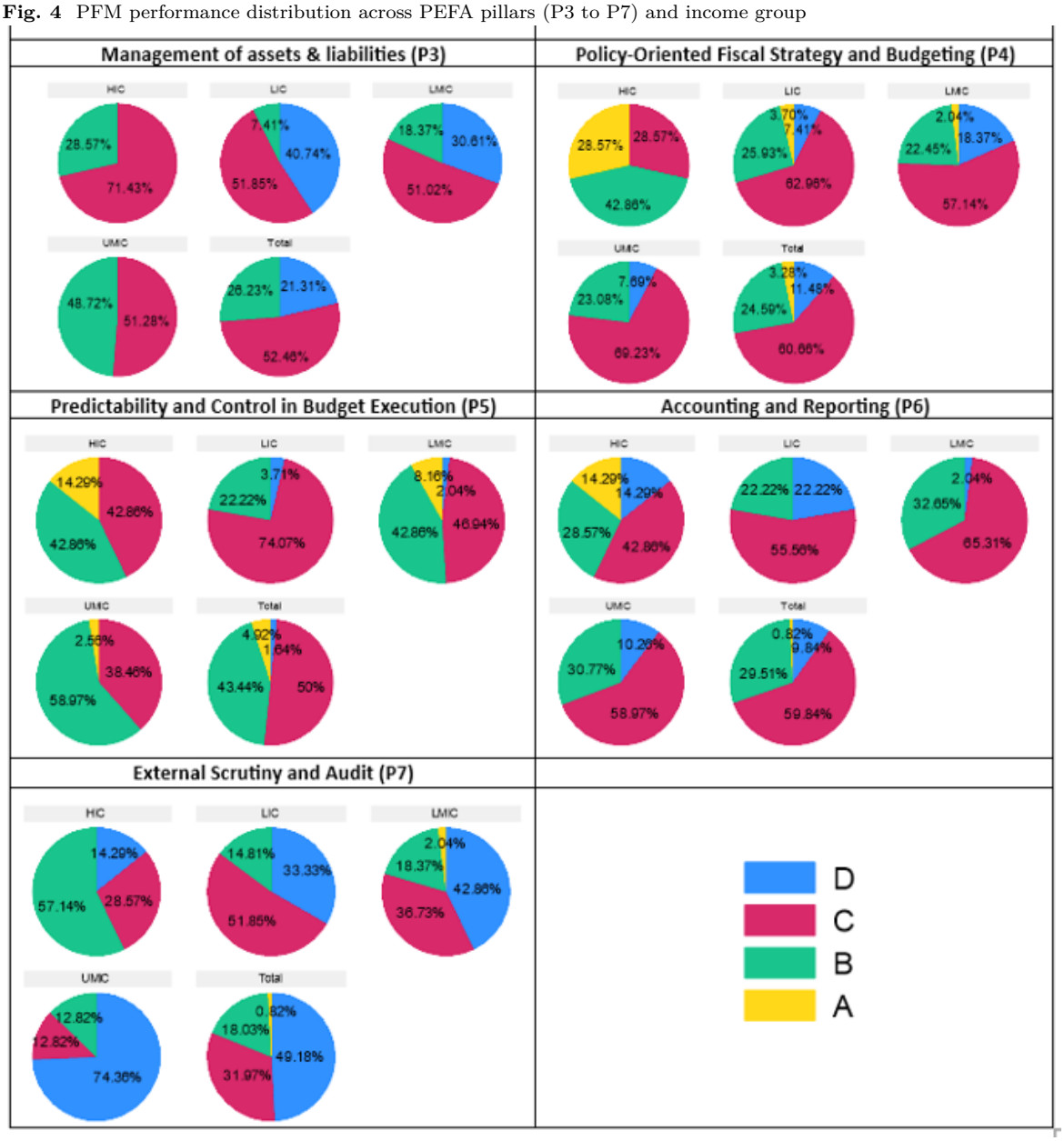
Note: A – Advanced, B – Better, C – Basic PFM practices as per international benchmark, D – Below basic.
Source: Data analysis

Fig. 3 Transparency of public finances (P2) performance distribution across income group



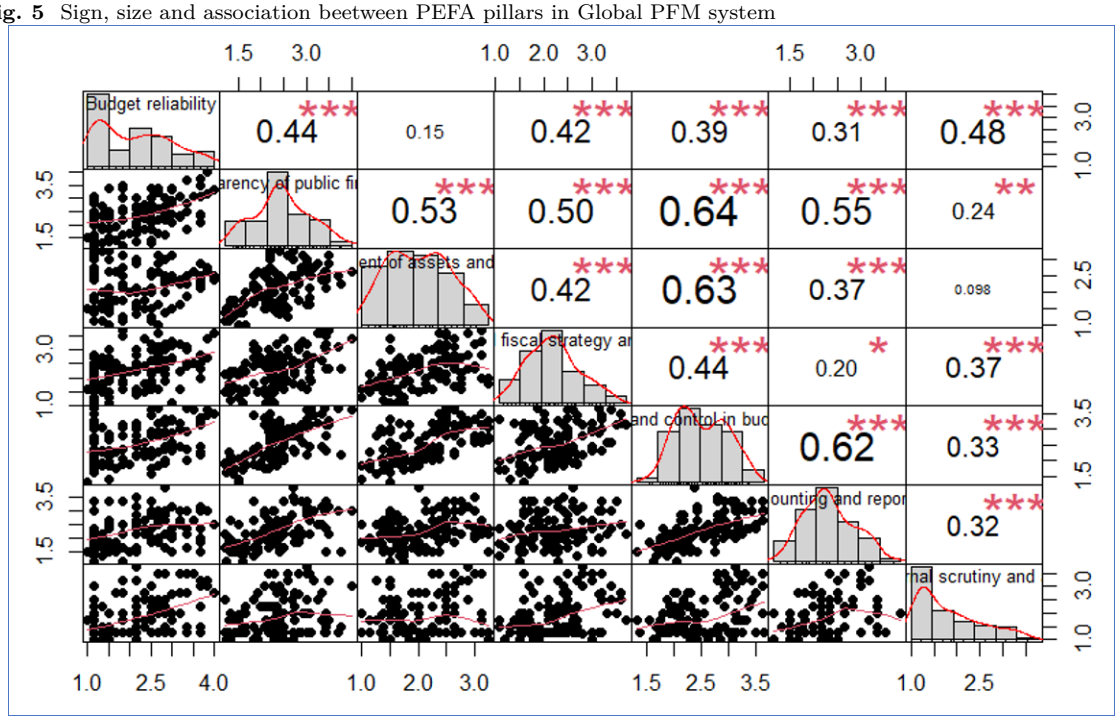
Note: A – Advanced, B – Better, C – Basic PFM practices as per international benchmark, D – Below basic.
Source: Data analysis

Fig. 4 PFM performance distribution across PEFA pillars (P3 to P7) and income group



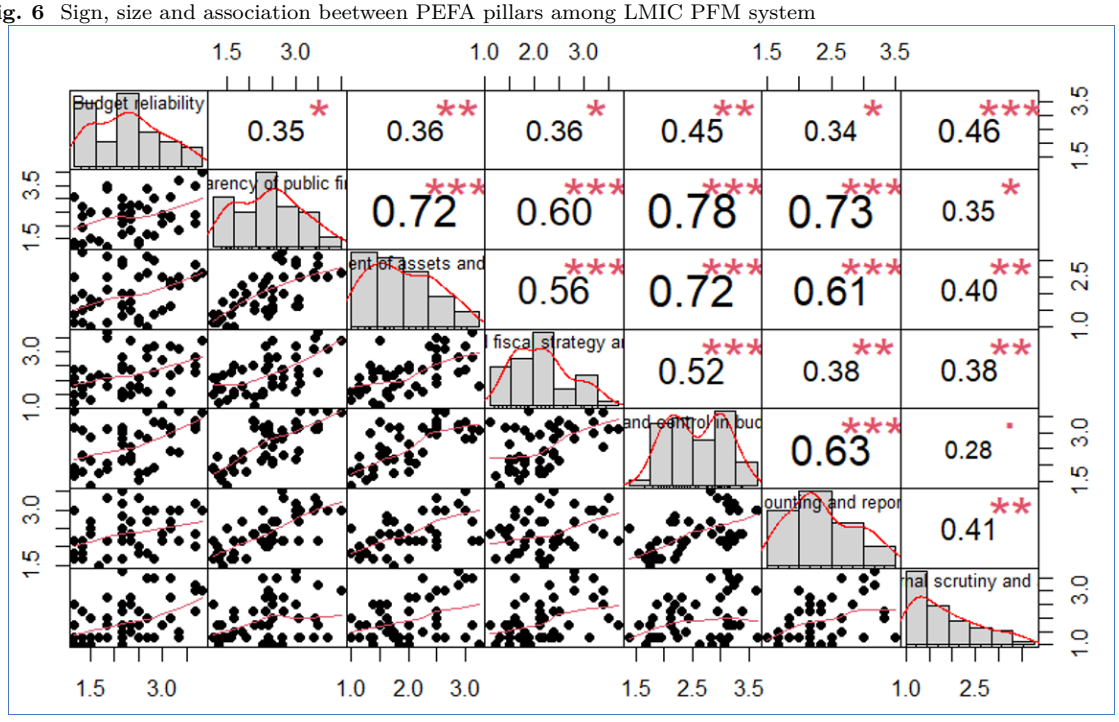
Note: A – Advanced, B – Better, C – Basic PFM practices as per international benchmark, D – Below basic.
Source: Data analysis

Fig. 5 Sign, size and association beetween PEFA pillars in Global PFM system



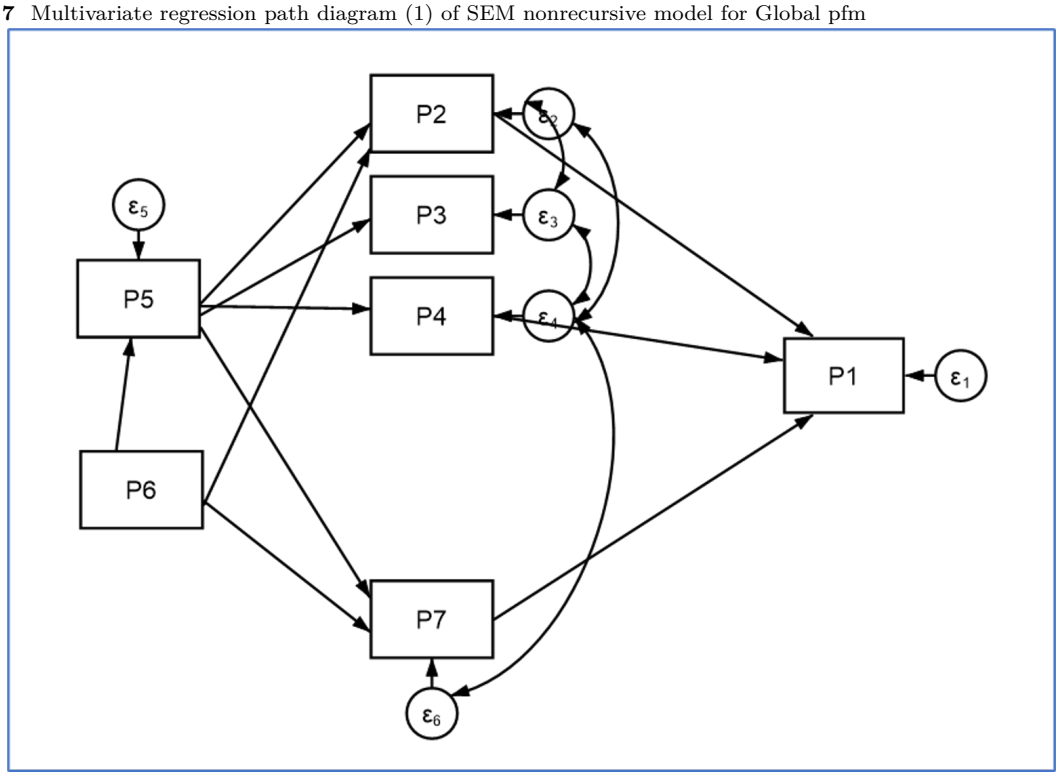
Note: ***, **, and * indicates significance level respectively at 1%, 5% and 10% error.
Source: Data analysis

Fig. 6 Sign, size and association beetween PEFA pillars among LMIC PFM system



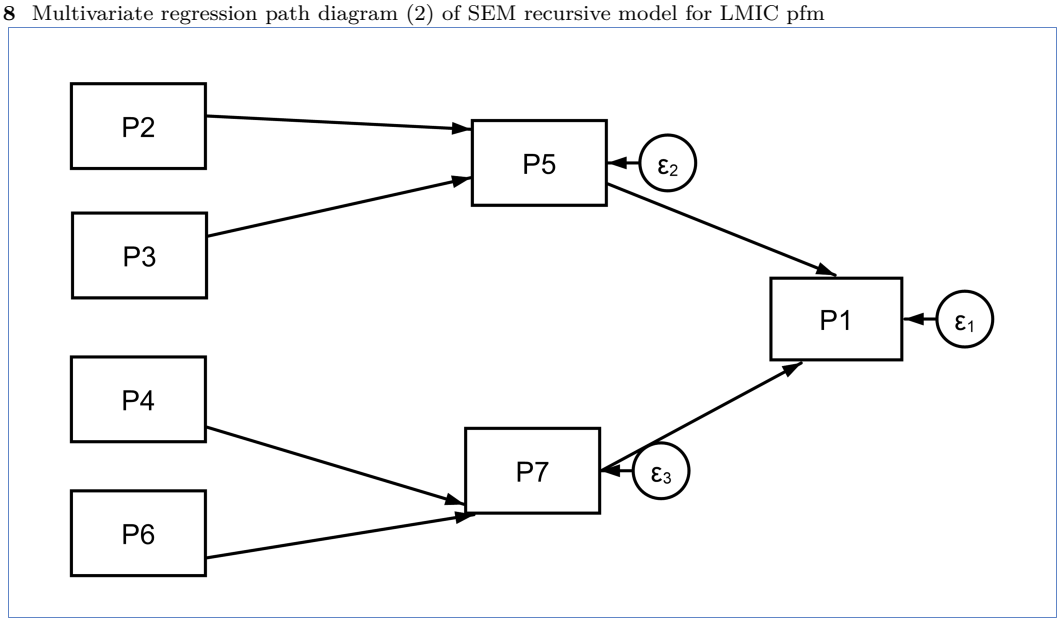
Note: ***, **, and * indicates significance level respectively at 1%, 5% and 10% error.
Source: Data analysis

Fig. 7 Multivariate regression path diagram (1) of SEM nonrecursive model for Global pfm



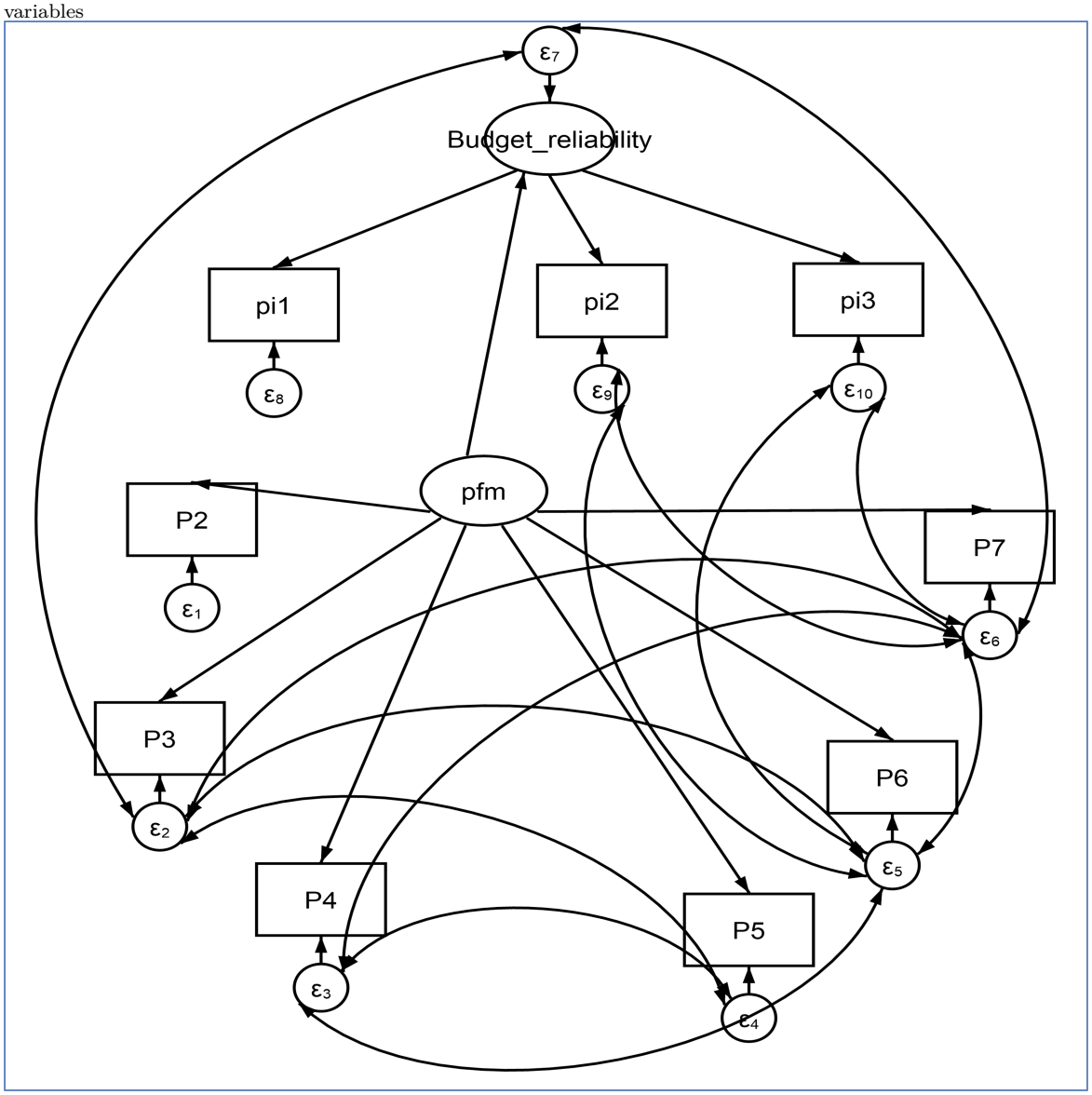
Note: P1 – Budget reliability; P2 – Transparency of public finances; P3 – Management of assets & liabilities; P4 – Policy-based fiscal strategy & budgeting; P5 – Predictability & control in budget execution; P6 – Accounting & reporting; P7 – External scrutiny & audit; $\varepsilon 1$ - $\varepsilon 6$ are error terms.
Source: Data analysis

Fig. 8 Multivariate regression path diagram (2) of SEM recursive model for LMIC pfm



Note: P1 – Budget reliability; P2 – Transparency of public finances; P3 – Management of assets & liabilities; P4 – Policy-based fiscal strategy & budgeting; P5 – Predictability & control in budget execution; P6 – Accounting & reporting; P7 – External scrutiny & audit; $\varepsilon 1$ - $\varepsilon 3$ are error terms.
Source: Data analysis

Fig. 9 Confirmatory factor analysis (CFA) path diagram (3) of equivalent measurement model introducing latent variables



Note: Latent variables are Budget-reliability, and pfm; Observed variables are: pi1 – Aggregate expenditure outturn; pi2 – Expenditure composition outturn; pi3 – Revenue outturn; P2 – Transparency of public finances; P3 – Management of assets & liabilities; P4 – Policy-based fiscal strategy & budgeting; P5 – Predictability & control in budget execution; P6 – Accounting & reporting; P7 – External scrutiny & audit; $\varepsilon 1 - \varepsilon 10$ are error terms.
Source: Data analysis

Table(s):

Table 2 GOF measures on evaluation of budget accountability (Adapted from Browne and Cudeck 1993; Hu and Bentler 1999; Kang & Ahn 2021)

Model fitted criteria		Recommended level of goodness of fit indicated by:	Structural model		Measurement model	
			Global	LMIC	Global	LMIC
1.	χ^2/df	1 to 2	1.247	0.387	1.569	0.939
2.	Prob > Chi-square_ms	p-value > 0.05 (good fit)	0.273	0.944	0.086	0.511
3.	Prob > Chi-square_bs	p-value > 0.05 (good fit)	0.000	0.000	0.000	0.000
4.	RMSEA	< 0.05 (very good) to 0.08 (adequate fit)	0.045	0.000	0.069	0.000
5.	SRMR	< 0.08	0.039	0.024	0.049	0.036
6.	CFI	0 (no fit) to 1 (perfect fit); > 0.95, sometimes 0.90 (good fit)	0.995	1.000	0.983	1.000
7.	TLI	0 (no fit) to 1 (perfect fit); > 0.95 (good fit)	0.984	1.126	0.952	1.1011
8.	AIC	Used for comparing alternate models:	1406.318	536.880	2049.758	848.277
9.	BIC	smaller (in absolute value) is better	1479.223	559.582	2164.723	925.842
10.	Overall SMC	Values closer to 1 indicate good fit	0.467	0.708	0.935	0.961
11.	Stability index	All the eigenvalues lie inside the unit circle. SEM satisfies stability condition.	0	0	0	0
12.	Wald-tests for equations	p-value < 0.05 (the endogenous variables are caused by the exogenous variables)	All equations satisfy (p-value = 0.000)	All equations satisfy (p-value = 0.000)	All equations satisfy (p-value = 0.000)	All equations satisfy (p-value = 0.000)

Note: df = degree of freedom; RMSEA = root mean squared error of approximation; SRMR = standardized root mean square residual; CFI = comparative normed fit index; TLI = Trucker-Lewis index; AIC = Akaike's information criterion; BIC = Bayesian information criterion; GFI = goodness-of-fit-index.

Source: Data analysis

Table 3 Loadings of measurement model for Global PFM

A. Estimates of loadings					
Latent variables	Measurement variables	Loadings	Standardized estimates	p-value	SMC
PFM system	Transparency of public finances (P2)	1.000	0.777***	–	0.604
	Management of assets & liabilities (P3)	0.812	0.676***	0.000	0.457
	Policy-based fiscal strategy & budgeting (P4)	0.798	0.659***	0.000	0.435
	Predictability & control in budget execution (P5)	0.955	0.811***	0.000	0.657
	Accounting & reporting (P6)	0.856	0.740***	0.000	0.547
	External scrutiny & audit (P7)	0.604	0.414***	0.000	0.171
Budget reliability	Aggregate expenditure outturn (PI-1)	1.000	0.758***	–	0.575
	Expenditure composition outturn (PI-2)	0.635	0.719***	0.000	0.549
	Revenue outturn (PI-3)	0.753	0.741***	0.000	0.390
B. Covariances of measurement error					
Management of assets & liabilities (P3)	Accounting & reporting (P6)	-0.041	-0.242*	0.062	
	External scrutiny & audit (P7)	-0.060	-0.205**	0.039	
	Budget reliability	-0.130	-0.401***	0.004	
Policy-based fiscal strategy & budgeting (P4)	Accounting & reporting (P6)	-0.089	-0.508***	0.000	
Accounting & reporting (P6)	Revenue outturn (PI-3)	-0.074	-0.299**	0.022	
External scrutiny & audit (P7)	Expenditure composition outturn (PI-2)	-0.091	-0.238*	0.049	
	Budget reliability	0.229	0.471***	0.001	

Note: ***, **, and * indicates significance level respectively at 1%, 5% and 10% error.

Source: Data analysis

Table 4 Estimates of Structural Model of latent variables

A. Regression weights					
Endogenous variables	Exploratory variables	Estimates	Standardized estimates	p-value	SMC
In case of global pfm					
Budget reliability (latent variable)	PFM (latent variable)	1.181	0.625***	0.000	0.390
In case of LMIC pfm					
Budget reliability (latent variable)	PFM (latent variable)	0.612	0.629***	0.004	0.396

Note: *** indicates significance level at 1% error.

Source: Data analysis

Table 5 Estimates of Structural Model for Global PFM

Endogenous variables	Exploratory variables	Estimates	Standardized estimates	p-value	SMC
A. Regression weights					
Budget reliability (P1)	Transparency of public finances (P2)	0.377	0.285***	0.001	0.366
	Policy-based fiscal strategy & budgeting (P4)	0.204	0.145*	0.098	
	External scrutiny & audit (P7)	0.418	0.360***	0.000	
Transparency of public finances (P2)	Predictability & control in budget execution (P5)	0.507	0.462***	0.000	0.450
	Accounting & reporting (P6)	0.311	0.279***	0.000	
Management of assets & liabilities (P3)	Predictability & control in budget execution (P5)	0.639	0.626***	0.000	0.392
Policy-based fiscal strategy & budgeting (P4)	Predictability & control in budget execution (P5)	0.450	0.436***	0.000	0.189
Predictability & control in budget execution (P5)	Accounting & reporting (P6)	0.630	0.619***	0.000	0.382
External scrutiny & audit (P7)	Predictability & control in budget execution (P5)	0.234	0.187*	0.074	0.138
	Accounting & reporting (P6)	0.288	0.226**	0.026	
B. Covariances of exogenous variables					
Transparency of public finances (P2)	Management of assets & liabilities (P3)	0.053	0.238***	0.010	
Transparency of public finances (P2)	Policy-based fiscal strategy & budgeting (P4)	0.092	0.356***	0.000	
Management of assets & liabilities (P3)	Policy-based fiscal strategy & budgeting (P4)	0.062	0.245***	0.006	
Policy-based fiscal strategy & budgeting (P4)	External scrutiny & audit (P7)	0.112	0.303***	0.001	

Note: ***, **, and * indicates significance level respectively at 1%, 5% and 10% error.

Source: Data analysis

Table 6 Estimates of Structural Model for LMIC PFM

A. Regression weights					
Endogenous variables	Exploratory variables	Estimates	Standardized estimates	p-value	SMC
Budget reliability (P1)	Predictability & control in budget execution (P5)	0.423	0.351***	0.004	0.334
	External scrutiny & audit (P7)	0.364	0.360***	0.003	
Predictability & control in budget execution (P5)	Transparency of public finances (P2)	0.465	0.534***	0.000	0.661
	Management of assets & liabilities (P3)	0.340	0.339***	0.005	
External scrutiny & audit (P7)	Policy-based fiscal strategy & budgeting (P4)	0.316	0.1269**	0.047	0.226
	Accounting & reporting (P6)	0.451	0.304**	0.025	

Note: ***, **, and * indicates significance level respectively at 1%, 5% and 10% error.
Source: Data analysis

Table 7 Mediating effects of structural model for Global PFM

Endogenous variables	Exploratory variables	Direct effects	Indirect effects	Total effects
Budget reliability (P1)	Transparency of public finances (P2)	0.377***	–	0.377***
	Policy-based fiscal strategy & budgeting (P4)	0.204*	–	0.204*
	Predictability & control in budget execution (P5)	–	0.381***	0.381***
	External scrutiny & audit (P7)	0.418***	–	0.418***
	Accounting & reporting (P6)	–	0.478***	0.478***
Transparency of public finances (P2)	Predictability & control in budget execution (P5)	0.507***	–	0.507***
	Accounting & reporting (P6)	0.311**	0.319***	0.631***
Management of assets & liabilities (P3)	Predictability & control in budget execution (P5)	0.639***	–	0.639***
	Accounting & reporting (P6)	–	0.402***	0.402***
Policy-based fiscal strategy & budgeting (P4)	Predictability & control in budget execution (P5)	0.450***	–	0.450***
	Accounting & reporting (P6)	–	0.283***	0.283***
Predictability & control in budget execution (P5)	Accounting & reporting (P6)	0.630***	–	0.630***
External scrutiny & audit (P7)	Predictability & control in budget execution (P5)	0.234*	–	0.234*
	Accounting & reporting (P6)	0.288**	0.147*	0.435***

Note: ***, **, and * indicates significance level respectively at 1%, 5% and 10% error.
Source: Data analysis

Table 8 Mediating effects of structural model for LMIC PFM

Endogenous variables	Exploratory variables	Direct effects	Indirect effects	Total effects
Budget reliability (P1)	Predictability & control in budget execution (P5)	0.423***	–	0.423***
	External scrutiny & audit (P7)	0.364***	–	0.364***
	Transparency of public finances (P2)	–	0.197**	0.197**
	Accounting & reporting (P6)	–	0.164*	0.164*
	Management of assets & liabilities (P3)	–	0.144**	0.144**
	Policy-based fiscal strategy & budgeting (P4)	–	0.115*	0.115*
Predictability & control in budget execution (P5)	Transparency of public finances (P2)	0.465***	–	0.465***
	Management of assets & liabilities (P3)	0.340***	–	0.340***
External scrutiny & audit (P7)	Policy-based fiscal strategy & budgeting (P4)	0.316**	–	0.316**
	Accounting & reporting (P6)	0.451**	–	0.451**

Note: ***, **, and * indicates significance level respectively at 1%, 5% and 10% error.

Source: Data analysis

Table 9 Power analysis for parameter estimated in SEMs (Adapted from Wang & Rhemtulla 2021)

Structural models						Measurement models		
Global (N=122)			LMIC (N=49)			(standardized)		
Parameter	Unstd.	Std.	Parameter	Unstd.	Std.	Parameter	Global	LMIC
P1∼P2	0.97	0.83	P1∼P5	0.92	0.86	P1 =∼pi1	1	0.89
P1∼P4	0.55	0.34	P1∼P7	0.81	***0.76***	P1 =∼pi2	1	0.98
P1∼P7	0.99	0.97	P5∼P2	0.84	0.95	P1 =∼pi3	1	0.94
P2∼P5	1	1	P5∼P3	0.67	0.68	pfm =∼P2	1	1
P2∼P6	0.95	0.91	P7∼P4	0.63	0.18	pfm =∼P3	1	1
P4∼P5	1	1	P7∼P6	0.88	0.56	pfm =∼P4	1	0.99
P7∼P5	***0.79***	0.56				pfm =∼P5	1	1
P7∼P6	0.91	***0.69***				pfm =∼P6	1	0.99
P3∼P5	1	1				pfm =∼P7	0.97	***0.69***
						P1∼pfm	1	***0.68***

Note: Unstd. = Unstandardized; Std. = Standardized; Latent variables are P1 (Budget-reliability), and pfm; Observed variables are: pi1 – Aggregate expenditure outturn; pi2 – Expenditure composition outturn; pi3 – Revenue outturn; P1 – Budget reliability; P2 – Transparency of public finances; P3 – Management of assets & liabilities; P4 – Policy-based fiscal strategy & budgeting; P5 – Predictability & control in budget execution; P6 – Accounting & reporting; P7 – External scrutiny & audit; =∼ "is measured by"; ∼ "is regressed on"; Alpha level is set at 0.05, when it is set at 10% significance level, then the power value increases such as: pfm =∼P7 becomes 0.79, and P1 ∼ pfm becomes 0.84, P7∼P6 (global, standardized) becomes 0.82, P7∼P5 (global, unstandardized) becomes 0.84, P1∼P7 (LMIC, standardized) becomes 0.81.

Source: Data analysis

MORE IN THE SERIES

- Kaur Amandeep and Lekha Chakraborty (2025). COP30 and Fiscal Policy Analysis for Climate Adaptation in an Emerging Economy, India, WP No. 434 (August).
- Mukherjee Sacchidananda, (2025). Reflecting on the Past to Plan the Future: Revenue Performance Evaluation of Indian GST, WP No. 433 (August).
- Badola Shivani and Sacchidananda Mukherjee, (2025). Beyond Necessities: A Closer Look at Per Capita Consumption Expenditure on Tobacco and Alcoholic Beverages during the last Decade WP No. 432 (July).

Dinesh Kumar Nayak, Economist, NIPFP, Email: dinesh.nayak@nipfp.org.in

Pratap Ranjan Jena, Professor, NIPFP, Email: pratap.jena@nipfp.org.in

National Institute of Public Finance and Policy,
18/2, Satsang Vihar Marg,
Special Institutional Area (Near JNU),
New Delhi 110067
Tel. No. 26569303, 26569780, 26569784
Fax: 91-11-26852548
www.nipfp.org.in