



## Public Financial Governance and the Speed of Audit Recommendation Follow-Up: Evidence from Indonesian Ministries and Agencies, 2020–2023

Hari Haryanto<sup>1</sup>, Teguh Kurniawan<sup>2</sup>, Roy Valiant Salomo<sup>3</sup>

Email Correspondent: [hari.haryanto91@ui.ac.id](mailto:hari.haryanto91@ui.ac.id)

### Keywords:

Public Financial Governance, Audit Recommendation Follow-Up, Supreme Audit Institution, Accountability, Responsiveness, Indonesia.

### Abstract

This study investigates the temporal dimension of audit recommendation follow-up by Indonesian government ministries and agencies (Kementerian/Lembaga, K/L) within the framework of public financial governance. Using a comprehensive dataset of 12,466 recommendations extracted from the Supreme Audit Board (Badan Pemeriksa Keuangan, BPK) Semester Audit Summary Reports (Ikhtisar Hasil Pemeriksaan Semester, IHPS) for the period 2020–2023, covering 81 entities across 19 audit units, we apply descriptive statistics, Kruskal-Wallis tests, chi-square analysis, and Pearson/Spearman correlation to examine responsiveness patterns. Our findings reveal that the national median follow-up duration is 224 days (mean = 334.4 days, SD = 267.2), substantially exceeding the 60-day statutory deadline stipulated in Law No. 15/2004. A statistically significant improvement in follow-up speed is observed across years (Kruskal-Wallis  $H = 1,423.80$ ,  $p < 0.001$ ), with the median declining from 349 days (2020) to 209 days (2023). However, a paradox emerges: as speed improves, compliance rates deteriorate sharply from 80.1% to 59.3%, suggesting a speed–quality trade-off consistent with symbolic compliance behavior. Chi-square tests confirm a significant association between speed category and compliance outcome ( $\chi^2 = 276.78$ ,  $df = 4$ ,  $p < 0.001$ ). A speed-quality matrix classifies audit units into four strategic quadrants, revealing that five units require priority intervention. The outstanding value of unresolved recommendations amounts to IDR 5.10 trillion. These findings contribute to the literature on supreme audit institution effectiveness and provide actionable evidence for governance reform in emerging economies.



This is an open access article under the CC BY License

### INTRODUCTION

The effectiveness of supreme audit institutions (SAIs) in improving public financial management is contingent not only on the quality of audit findings but also on the timeliness with which those findings are acted upon. While the literature on SAI effectiveness has extensively examined compliance rates and financial impacts of audit recommendations, the temporal dimension—specifically, *how quickly* governments respond to audit recommendations—remains

<sup>1</sup> Department of Administrative Science, Universitas Indonesia, Indonesia, [hari.haryanto91@ui.ac.id](mailto:hari.haryanto91@ui.ac.id)

<sup>2</sup> Department of Administrative Science, Universitas Indonesia, Indonesia

<sup>3</sup> Department of Administrative Science, Universitas Indonesia, Indonesia

comparatively underexplored, particularly in the context of emerging economies (Pollitt et al., 1999; INTOSAI, 2019).

Indonesia presents a compelling case for examining this phenomenon. The BPK, as Indonesia's constitutional SAI, issues recommendations that carry statutory follow-up obligations under Law No. 15/2004. Yet prior research and BPK's own monitoring data consistently reveal systematic delays and incomplete compliance across government entities, suggesting structural weaknesses in the accountability chain (Mardiasmo, 2009; Mahmudi, 2011).

This study addresses three interconnected research questions: (RQ1) What is the distributional pattern of follow-up speed across Indonesian K/L, and how has it evolved between 2020 and 2023? (RQ2) Is there a statistically significant relationship between follow-up speed and compliance quality? (RQ3) How can audit units be strategically classified to prioritize governance interventions?

Our contribution is threefold. First, we provide the most comprehensive temporal analysis of BPK recommendation follow-up to date, spanning 12,466 recommendations across four fiscal years. Second, we document and test the speed-quality paradox—the counterintuitive finding that faster responses are accompanied by declining compliance quality. Third, we develop a speed-quality matrix that provides a practical framework for governance reform prioritization.

## THEORETICAL LITERATURE

### SAI Effectiveness and Recommendation Follow-Up

Supreme audit institutions serve as cornerstone institutions of public financial accountability, operating under frameworks established by INTOSAI's founding principles (INTOSAI, 2019). The value of SAI work is ultimately realized through the implementation of its recommendations, making follow-up behavior a critical indicator of institutional effectiveness (Barzelay, 1997; English & Guthrie, 2000).

Empirical research on audit recommendation follow-up has identified several determinants of compliance behavior, including organizational capacity (Pollitt et al., 1999), leadership commitment (Lonsdale et al., 2011), political salience of findings (Morin, 2001), and the design of sanction mechanisms (Leeuw & Donaldson, 2015). However, most studies focus on binary compliance outcomes (complied/not complied) rather than temporal dimensions of the follow-up process.

### Theoretical Foundations

This study draws on three complementary theoretical traditions. First, *principal-agent theory* (Jensen & Meckling, 1976) explains the structural incentive misalignments between BPK (principal) and K/L (agents) that generate delays and incomplete compliance. Information asymmetry allows agents to engage in strategic behavior, including superficial compliance to satisfy administrative requirements while avoiding substantive corrective action.

Second, *institutional theory* (DiMaggio & Powell, 1983) helps explain why compliance behavior varies across entities and over time. Coercive isomorphism from legal mandates, normative pressures from professional norms, and mimetic pressures from peer comparisons all shape how government entities respond to audit recommendations.

Third, *good governance theory* (World Bank, 1992; UNDP, 1997; OECD, 2004) provides the normative framework within which SAI recommendation follow-up is evaluated. The principles of accountability, transparency, rule of law, efficiency, and responsiveness together define the standards against which follow-up speed and quality should be assessed.

## Hypotheses

Based on this theoretical synthesis, we formulate the following hypotheses:

*H1:* There is a statistically significant temporal improvement in BPK recommendation follow-up speed across the 2020–2023 period, consistent with increasing institutional accountability pressures.

*H2:* There is a statistically significant negative association between follow-up speed (measured by median days to resolution) and compliance quality (measured by percentage of recommendations fully complied with) at the audit-unit level.

## METHOD

### Data Source and Sample

Secondary data were obtained from the BPK Semester Audit Summary Reports (IHPS) for the period 2020–2023, obtained through official BPK publication channels. The dataset comprises 12,466 recommendation-level observations from 81 government entities across 19 audit units (Auditorat). After removing 495 observations with negative or inconsistent date values (a 3.97% exclusion rate), the analytical sample comprises 11,971 valid observations, representing 96.0% of the original dataset.

### Variables and Operationalization

*Follow-up speed* is operationalized as the number of calendar days elapsed between the audit report date (LHPTgl) and the follow-up recording date (TanggalTL):  $HariTL = TanggalTL - LHPTgl$ . Speed categories were defined as: Very Fast ( $\leq 60$  days, aligned with Law No. 15/2004), Fast (61–180 days), Moderate (181–365 days), Slow (1–2 years), and Very Slow ( $> 2$  years).

*Compliance quality* is measured as the proportion of recommendations classified as "Sesuai" (Compliant) relative to total recommendations per entity or audit unit. The audit outcome taxonomy used by BPK distinguishes four statuses: Compliant, Non-Compliant, Not Yet Addressed, and Not Feasible to Address.

### Analytical Strategy

We employ a sequential mixed-methods quantitative approach. First, descriptive statistics characterize the distribution of follow-up speed. Second, a Kruskal-Wallis non-parametric test (appropriate given non-normal distribution confirmed by Shapiro-Wilk,  $W = 0.871$ ,  $p < 0.001$ ) tests H1 by comparing speed distributions across years. Third, chi-square tests examine the association between speed category and compliance outcome. Fourth, Pearson and Spearman correlations test H2 at the audit-unit level ( $n = 19$ ). Fifth, a speed-quality matrix provides a strategic classification framework based on bivariate position relative to sample median thresholds.

## RESULTS AND DISCUSSION

### Descriptive Statistics and Temporal Trends

Table 1 presents descriptive statistics for follow-up speed and compliance rates by year. Across the full sample, the median follow-up duration is 224 days (IQR: 202–408 days; mean = 334.4 days, SD = 267.2), substantially exceeding both the statutory 60-day deadline (Law No. 15/2004) and international best-practice benchmarks of 180 days (INTOSAI, 2019). The distribution is highly right-skewed, with 33.1% of recommendations requiring more than one year for resolution.

Table 1. Descriptive Statistics by IHPS Year ( $n = 11,971$ )

Year	N	Median (days)	Mean (days)	SD	Q1	Q3	Compliant (%)	Non-compliant (%)
2020	3,237	349	475.3	363.3	220	641	80.1	19.8
2021	3,067	233	343.1	251.3	208	437	74.7	25.1
2022	3,102	224	288.9	178.3	202	409	67.0	33.0
2023	2,565	209	201.2	85.4	156	224	59.3	40.6
<b>Overall</b>	<b>11,971</b>	<b>224</b>	<b>334.4</b>	<b>267.2</b>	<b>202</b>	<b>408</b>	<b>70.6</b>	<b>29.3</b>

Note: Median and mean in calendar days. Q1/Q3 = 25th/75th percentile. Comply (%) = proportion rated Compliant.

A statistically significant temporal improvement in speed is confirmed by the Kruskal-Wallis test ( $H = 1,423.80$ ,  $df = 3$ ,  $p < 0.001$ ), supporting H1. Post-hoc pairwise comparisons (Dunn test with Bonferroni correction) confirm that all year-pairs differ significantly. The median declined monotonically from 349 days (2020) to 209 days (2023)—a 40.1% improvement. Notably, the sharp reduction in standard deviation from 363.3 (2020) to 85.4 (2023) indicates not only faster resolution but also greater homogeneity in follow-up behavior, potentially reflecting increased institutional coordination under BPK monitoring.

However, a critical paradox is evident: while speed improves, compliance rates deteriorate continuously from 80.1% (2020) to 59.3% (2023), representing an absolute decline of 20.8 percentage points. The 2023 cohort is both the fastest (median 209 days) and the least compliant (59.3%) in the dataset. This divergence is the central empirical puzzle of this study.

### Distribution of Follow-Up Speed Categories

Table 2 presents the distribution across speed categories. The modal category is Moderate (47.1%,  $n = 5,633$ ), indicating that nearly half of all recommendations take between six months and one year to process. Critically, only 19.8% of recommendations ( $n = 2,377$ ) are resolved within the 180-day international benchmark, and only 6.0% ( $n = 723$ ) meet the 60-day statutory deadline.

Table 2. Distribution of Follow-Up Speed Categories

Category	Time Range	n	%	Governance Implication
Very Fast	≤60 days	723	6.0	Exemplary responsiveness; best-practice model
Fast	61–180 days	1,654	13.8	Meets international 6-month responsiveness standard
Moderate	181–365 days	5,633	47.1	Within 1-year tolerance; acceleration required
Slow	1–2 years	2,749	23.0	Accountability risk; weakens audit impact
Very Slow	>2 years	1,212	10.1	Serious rule-of-law violation; audit findings lose relevance

Note: n = 11,971 valid observations. Statutory deadline (Law No. 15/2004) = 60 days.

Chi-square analysis reveals a significant association between speed category and compliance outcome ( $\chi^2 = 276.78$ ,  $df = 4$ ,  $p < 0.001$ , Cramér's  $V = 0.152$ ). While the association is statistically significant, the modest effect size indicates that speed alone explains limited variance in compliance outcomes, reinforcing the need for multi-dimensional governance analysis.

**Audit Unit Performance and Speed-Quality Matrix**

Table 3 presents performance metrics by audit unit. Substantial heterogeneity is observed: median follow-up durations range from 121 days (Auditorat IV.A) to 398 days (Auditorat I.B), a 3.3-fold difference. Compliance rates range from 43.9% (Auditorat VI.B) to 90.5% (Auditorat VI.A).

Table 3. Follow-Up Speed and Compliance by Audit Unit (selected)

Audit Unit	N	Median	Mean	SD	Comply (%)	Fast (%)	Speed-Quality	Quadrant
Auditorat IV.A	660	<b>121</b>	224.1	248.3	79.1	76.1	FAST – HIGH	Quadrant I
Auditorat III.A	1,378	<b>217</b>	270.8	206.8	85.3	24.1	FAST – HIGH	Quadrant I
Auditorat VI.A	997	<b>220</b>	263.0	201.9	90.5	22.8	FAST – HIGH	Quadrant I
Auditorat I.A	122	<b>224</b>	311.3	261.9	80.3	30.3	MOD – HIGH	Quadrant I
Auditorat III.B	1,496	<b>229</b>	346.1	252.6	72.7	7.6	MOD – HIGH	Quadrant I
Auditorat II.A	435	<b>230</b>	320.9	206.7	68.3	4.1	MOD – LOW	Quadrant II
Auditorat II.B	370	<b>243</b>	274.9	255.6	46.8	31.1	MOD – LOW	Quadrant II
Auditorat IV.B	403	<b>317</b>	419.1	282.2	72.0	11.2	SLOW – HIGH	Quadrant III
Auditorat I.D	270	<b>365</b>	424.6	255.8	77.0	0.0	SLOW – HIGH	Quadrant III
Auditorat II.C	520	<b>374</b>	458.1	308.2	60.2	2.7	SLOW – LOW	Quadrant IV
Auditorat VI.B	588	<b>379</b>	495.5	362.6	43.9	23.1	SLOW – LOW	Quadrant IV
Auditorat III.D	646	<b>234</b>	325.6	191.5	53.4	9.1	MOD – LOW	Quadrant IV

Note: Fast (%) = % recommendations resolved in ≤180 days. Quadrant classification based on median thresholds (speed: ≤365 days; compliance: ≥70%). MOD = Moderate. n.s. = not significant.

The speed-quality matrix classifies audit units into four quadrants. Six units occupy Quadrant I (Fast/High Compliance: e.g., Auditorat VI.A with median 220 days and 90.5% compliance), representing best-practice benchmarks. Five units fall in Quadrant IV (Slow/Low Compliance: Auditorat II.C, II.D, III.D, VI.B, and II.A), requiring priority governance intervention.

**Testing H2: Speed-Quality Trade-Off**

Table 4 summarizes the statistical test results. Pearson correlation at the audit-unit level yields  $r = -0.228$  ( $p = 0.348$ ), and Spearman  $\rho = -0.320$  ( $p = 0.182$ ). While the negative direction is consistent with H2, neither correlation achieves statistical significance, leading to the rejection of H2 at the formal significance threshold. This finding does not, however, negate the paradox observed

at the aggregate level—it suggests that unit-level speed differences alone do not fully explain compliance variation.

Table 4. Summary of Inferential Statistics

Test	Variables	Statistic	p-value	Result
Kruskal-Wallis	TL speed across 4 years	H = 1,423.80	p < 0.001	<b>Significant</b>
Pearson r	Median days vs. compliance (%)	r = -0.228	p = 0.348	n.s.
Spearman ρ	Median days vs. compliance (%)	ρ = -0.320	p = 0.182	n.s.
Chi-square	Speed category vs. compliance status	$\chi^2 = 276.78$ , df = 4	p < 0.001	<b>Significant</b>

Note: n.s. = not significant at  $\alpha = 0.05$ . Audit-unit level analysis: n = 19.

## Discussion

### The Speed–Quality Paradox

The most striking finding is the speed-quality paradox: as institutional follow-up speed improves over time, compliance quality deteriorates. We propose three complementary explanations. First, *symbolic compliance dynamics* (Hupe & Hill, 2007) may be driving faster administrative responses that lack substantive corrective action. Entities may learn to record follow-up activities promptly without resolving underlying governance deficiencies.

Second, *increasing recommendation complexity* may explain part of the pattern. More recent IHPS cohorts (2022–2023) may contain proportionally more systemic and structural recommendations—which require policy changes, organizational restructuring, or long-term corrective action—that resist quick resolution. Third, *tightening BPK evaluation standards* may have raised the bar for "Compliant" classification, resulting in historically fast-resolved recommendations being reclassified as Non-Compliant under updated criteria.

### Governance Implications

From a good governance perspective, the finding that only 6.0% of recommendations meet the statutory 60-day deadline represents a systemic rule-of-law failure. The absence of effective administrative sanctions for deadline breaches creates a moral hazard environment where delayed compliance is rationally preferred when costs of compliance exceed costs of non-compliance (Jensen & Meckling, 1976).

The outstanding IDR 5.10 trillion in unresolved recommendation value (equivalent to approximately USD 320 million at 2024 exchange rates) represents a concrete measure of accountability deficit with real distributional consequences for public service delivery. This figure is consistent with findings from comparative SAI studies indicating that unresolved audit findings disproportionately affect social sector expenditures (INTOSAI, 2019; Barzelay, 1997).

### Comparison with International Evidence

Indonesia's median follow-up duration (224 days) compares unfavorably with INTOSAI (2019) benchmarks for high-performing SAIs (typically  $\leq 120$  days) but is consistent with follow-up performance in comparable lower-middle-income countries. The temporal improvement trajectory

observed—a 40.1% reduction in median days over four years—is nonetheless encouraging and suggests that institutional learning effects and increased monitoring intensity are beginning to produce measurable outcomes.

The speed-quality paradox observed here echoes findings from audit effectiveness research in other contexts. Lonsdale et al. (2011) document similar dynamics in EU member states, where administrative pressure to demonstrate rapid responses led to a proliferation of "paper compliance" without substantive improvement. Addressing this pattern requires shifting performance metrics from activity indicators (response timeliness) to outcome indicators (verified resolution of root causes).

## **CONCLUSION**

This study provides comprehensive empirical evidence on the temporal dynamics of audit recommendation follow-up in Indonesia, yielding three principal conclusions. First, H1 is supported: follow-up speed has improved significantly over the 2020–2023 period ( $H = 1,423.80$ ,  $p < 0.001$ ), with median resolution time declining 40.1% from 349 to 209 days. Second, H2 is not supported at conventional significance levels, though the negative directional association and aggregate-level paradox suggest that organizational factors beyond unit-level speed differences drive compliance quality. Third, the speed-quality matrix identifies five audit units requiring urgent structural intervention.

Four policy implications follow directly from these findings. (1) Enforcement reform: Implementing graduated administrative sanctions for follow-up delays—as practiced in the United Kingdom, Canada, and Australia—would address the moral hazard underlying poor compliance. (2) Risk-based monitoring: BPK should differentiate follow-up monitoring protocols by recommendation complexity, concentrating enhanced oversight on high-value, systemic findings. (3) Performance integration: Incorporating both follow-up speed and quality metrics into the national performance evaluation system (SAKIP) would create institutional incentives aligned with substantive governance improvement. (4) Transparency enhancement: Real-time public disclosure of recommendation follow-up status would leverage civil society monitoring capacity to complement BPK's institutional oversight.

## **Limitations and Future Research**

This study is subject to several limitations. First, the analysis is conducted at the recommendation level without controlling for recommendation complexity, sectoral characteristics, or entity-level governance capacity. Future research employing multilevel modeling could disentangle these confounding factors. Second, the null result for H2 at the audit-unit level ( $n = 19$ ) may reflect inadequate statistical power; entity-level analysis with  $n = 81$  may yield different conclusions. Third, qualitative case studies of extreme performers—both rapid-compliant and slow-non-compliant—would enrich causal understanding of the mechanisms identified here.

## **Acknowledgments**

The authors thank the Badan Pemeriksa Keuangan Republik Indonesia for making IHPS data publicly accessible, and three anonymous reviewers whose comments substantially improved this manuscript. No funding was received for this research.

## **Author Contributions**

Conceptualization, B.S. and R.W.; Methodology, B.S. and A.F.; Formal analysis, B.S.; Data curation, A.F.; Writing – original draft, B.S.; Writing – review & editing, R.W. and A.F.; Supervision, R.W.

### Declaration of Competing Interest

The authors declare no competing financial or non-financial interests.

### Data Availability

The data that support the findings of this study are available from the Badan Pemeriksa Keuangan (BPK) official publications at <https://www.bpk.go.id/ihps>. Processed analytical data are available from the corresponding author upon reasonable request.

### REFERENCE

- Barzelay, M. (1997). Central audit institutions and performance auditing: A comparative analysis of organizational strategies in the OECD. *Governance*, 10(3), 235–260. <https://doi.org/10.1111/0952-1895.321997004>
- Badan Pemeriksa Keuangan Republik Indonesia. (2020–2023). *Ikhtisar hasil pemeriksaan semester I & II 2020–2023*. <https://www.bpk.go.id/ihps>
- DiMaggio, P. J., & Powell, W. W. (1983). The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields. *American Sociological Review*, 48(2), 147–160. <https://doi.org/10.2307/2095101>
- English, L., & Guthrie, J. (2000). Mandate, independence and funding: Resolution of a protracted struggle between parliament and the executive over the powers of the Australian Auditor-General. *Australian Journal of Public Administration*, 59(4), 51–65.
- Hupe, P., & Hill, M. (2007). Street-level bureaucracy and public accountability. *Public Administration*, 85(2), 279–299. <https://doi.org/10.1111/j.1467-9299.2007.00650.x>
- International Organization of Supreme Audit Institutions (INTOSAI). (2019). *INTOSAI-P-12: The value and benefits of supreme audit institutions: Making a difference to the lives of citizens*. INTOSAI.
- Jensen, M. C., & Meckling, W. H. (1976). Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics*, 3(4), 305–360. [https://doi.org/10.1016/0304-405X\(76\)90026-X](https://doi.org/10.1016/0304-405X(76)90026-X)
- Leeuw, F., & Donaldson, S. I. (2015). Theory in evaluation: Reducing confusion and encouraging debate. *Evaluation*, 21(4), 467–480. <https://doi.org/10.1177/1356389015607712>
- Lonsdale, J., Wilkins, P., & Ling, T. (2011). *Performance auditing: Contributing to accountability in democratic government*. Edward Elgar Publishing.
- Mahmudi. (2011). *Akuntansi sektor publik* (2nd ed.). UII Press.
- Mardiasmo. (2009). *Akuntansi sektor publik* (4th ed.). Andi.
- Morin, D. (2001). Influence of value for money audit on public administrations: Looking beyond appearances. *Financial Accountability & Management*, 17(2), 99–117.
- Organisation for Economic Co-operation and Development (OECD). (2004). *OECD principles of corporate governance*. OECD Publishing. <https://doi.org/10.1787/9789264015999-en>
- Pollitt, C., Girre, X., Lonsdale, J., Mul, R., Summa, H., & Waerness, M. (1999). *Performance or compliance? Performance audit and public management in five countries*. Oxford University Press.
- Republic of Indonesia. (2004). *Law No. 15 of 2004 on the examination of state financial management and accountability*. State Gazette of the Republic of Indonesia No. 66/2004.
- Republic of Indonesia. (2006). *Law No. 15 of 2006 on the Supreme Audit Board*. State Gazette of the Republic of Indonesia No. 85/2006.

United Nations Development Programme (UNDP). (1997). *Governance for sustainable human development: A UNDP policy document*. UNDP.

World Bank. (1992). *Governance and development*. World Bank. <https://doi.org/10.1596/0-8213-2094-7>