eISSN 2656-8209 | pISSN 2656-1565
JURNAL INOVASI ILMU SOSIAL DAN POLITIK (JISoP)
Vol. 8, No. 2 (2026) pp. 131-140
Accreditation S3 – Decree No. 177/E/KPT/2024
https://riset.unisma.ac.id/index.php/JISoP/article/view/25233
DOI: 10.33474/jisop.v8i2.25233

Leveraging the Dalan Ba Futuru system to enhance the relationship between financial accountability and institutional performance in Timor-Leste

Mariano Nunes[1]*, and **Rudy Badrudin[2]**
[1]Timorese Catholic University, Díli, Timor-Leste
[2]Sekolah Tinggi Ilmu Ekonomi Yayasan Keluarga Pahlawan Negara, Yogyakarta, Indonesia
*email corresponding author: neonunes4@gmail.com

Article Info

Article history:
Submission: **2026-03-09**
Accepted: **2026-06-19**
Published: **2026-07-11**

ABSTRACT

In Timor-Leste, the shift to program-based budgeting requires robust accountability mechanisms to ensure public funds effectively enhance institutional performance. However, there is a lack of evidence regarding how accountability- specifically in planning, monitoring, and evaluating-relates to performance outcomes within the local context. Furthermore, the role of the Dalan Ba Futuru System (DBFTL), a web-based management tool, in moderating these processes remains under-researched. Objective: this study examines the influence of accountability in planning, monitoring and evaluation on the quality of state institution performance in Timor-Leste, and analyzes whether the DBFTL moderates theses relationship. Methodology: this study adopts a qualitative research design, utilizing primary data collected through questionnaires from 157 respondents, including government directors and technical staff involved in financial management. Data were analyzed using Partial Least Squares Equation Modeling (PLS-SEM). Findings: Empirical results indicate that accountability in monitoring has a significant positive effect on institutional performance (β= 0.346, p = 0.004), whereas planning and evaluation accountability show positive but insignificant effect. Notably, the DBFTL significantly moderates and strengthens the relationships between accountability practices and institutional performance ((β = 0.582, p < 0.001). Implications: The findings demonstrate that digital information system are critical for public sector reform, providing empirical support for the government to prioritize the integration of DBFTL System to enhance transparency and budget effectiveness.

Keywords:
financial accounting; public financial management; institutional performance

Please cite this article in APA style as:
Nunes, M., & Badrudin, R. (2026). Leveraging the Dalan Ba Futuru system to enhance the relationship between financial accountability and institutional performance in Timor-Leste. *Jurnal Inovasi Ilmu Sosial dan Politik (JISoP), 8*(2), 131–140. https://doi.org/10.33474/jisop.v8i2.25233

INTRODUCTION

Public financial accountability is the cornerstone of effective governance, ensuring that public resources are managed transparently to meet the needs and welfare of society (Alkaraan, 2018; Birdayanthi et al., 2025). High levels of accountability are linked to improved public trust and the efficient delivery of government services, which are critical indicators of institutional performance. In line with international standards, Timor-Leste has shifted its budgeting model toward performance-based budgeting to better evaluate the impact of public

spending. This reform is facilitated by the National Agency for Planning, Monitoring, and Evaluation (ANAPMA), which acts as a watchdog to ensure that line ministries report key results in accordance with established legal frameworks. To support these reforms, the government implemented the Dalan Ba Futuru System (DBFTL), a web-based information system designed to integrate planning, monitoring, and evaluation. By providing a structured, digital platform, the DBFTL aims to improve data accessibility and budget tracking, serving as a critical tool for public administration. While scholars like Cuadrado-ballesteros & Bisogno (2022) have explored the relationship between transparency and financial sustainability, and Miranda et al. (2008) have emphasized the need for technology-driven budget information, existing research often overlooks the unique regulatory and cultural context of Timor-Leste. Other studies (Pramudiati & Siregar, 2018; Soares & Wiagustini, 2019) have analyzed regional expenditures and accrual accounting but have yet to examine how specific technological tools like the DBFTL moderate the influence of accountability on performance. Research by (Fatmawati et al., 2026) confirms that incumbent politicians tend to manipulate public budgets before elections to increase their chances of winning. Therefore, the existence of an accountability system such as Dalan Ba Futuru is crucial as a monitoring tool to mitigate such opportunistic behavior. By strengthening anti-corruption controls and information transparency, this system is able to maintain fiscal integrity and strengthen institutional performance amid a political cycle that is prone to budget misuse. Consequently, there is a clear empirical gap regarding how digital management systems can strengthen accountability outcomes in emerging public sectors. This study is grounded in Principal-Agent Theory, which explains the necessity of accountability mechanisms when information asymmetry exists between the government and the public (Broadbent et al., 1996; Nyman et al., 2005), while further incorporating Information Systems Success concepts to justify the moderating role of the DBFTL, positing that a well-integrated system amplifies the effectiveness of human-led planning and monitoring activities. Therefore, this study aims to examine the influence of accountability in the planning, monitoring, and evaluation of the Timor-Leste State Budget on the quality of performance of state institutions, and to analyze the moderating role of the Dalan Ba Futuru System (DBFTL) in this relationship

The budgeting model has now shifted to program-based budgeting with more emphasis on performance budgeting. The purpose of this change is to encourage the evaluation of public spending based on the impact and benefits it provides to society, as well as to improve the overall accountability of public administration. This state body has an important role in facilitating the change in budgeting model and ensuring that the programs implemented by all line ministries (*Linhas Ministério*) are in line with established objectives. Through the quarterly and annual performance reports prepared by the *Agência Nacional De Planeamento Monitorização e Avaliação* (in English, the National Agency for Planning, Monitoring and Evaluation) or ANAPMA for short. This state body acts as a watchdog and evaluator to ensure the achievement of key results of these programs and to ensure accountability from the public administration to society. The performance reports produced by ANAPMA highlight the key results achieved by the programs over a given period. This is done using indicators and targets that have been set based on the goals and objectives of each program. This performance report reflects the changes in the presentation format of the performance report since the enactment of Decree Law No. 1 of 2023 at the beginning of January 2023 (Timor-Leste Ministry of Finance, 2024).

Based on the performance report prepared by ANAPMA, this State Agency acts as the entity that facilitates the reporting of performance progress of Public Administration Sector agencies and services or commonly called Line Ministries. Through DBFTL's planning, monitoring and evaluation information system, the State Agency ensures that the reports include relevant information and comply with the legal provisions applicable in Timor-Leste. The main objective is to measure the achievement of key results of these programs and improve the accountability and transparency of the public administration to the community (Timor-Leste Ministry of Finance, 2024).

National Agency for Planning, Monitoring and Evaluation (ANAPMA). This body is responsible for the submission of quarterly and annual performance reports of each of the programs implemented by the Public Administration Sector, in accordance with the new reporting rules established by subsection 3 of article 66 of Law No. 2 of 2022, dated 10 February. ANAPMA has the responsibility to organize, coordinate, and supervise the administration and implementation of multi-year government program planning: “The national medium- and long-term programs of the 8th Constitutional Government and the incoming Constitutional Government. This State Agency acts as the entity that facilitates the reporting of performance progress of Public Administration Sector agencies and services or commonly referred to as line ministries. Through DBFTL's planning, monitoring and evaluation information system, the State Agency ensures that the reports include relevant information and comply with the legal provisions applicable in Timor-Leste. The main objective is to measure the achievement of key results of these programs and improve the accountability and transparency of public administration to the community (Timor-Leste Ministry of Finance, 2024).

DBFTL is a tool or instrument used to develop annual plans, monitor and track their implementation, and evaluate their results to assess their achievements. One of the important aspects of DBFTL is the ability to show the changes and effects from Outcome to other results based on the sequence of Inputs, with the ultimate goal of achieving the set objectives and must comply with the provisions of the Budget Law applicable in the Government of Timor-Leste. DBFTL is also a web-based application developed with "Open Source" technology, so there are no license fees charged for users.

METHOD

This study included 48 active civil servants and regional council secretariats as the study population, based on data from the Civil Service Commission (CFP) in 2023. This group consists of Director Generals, National Directors, and technical staff who have authority under the Department of the Directorate General of Planning and Budget (DGPO). They are responsible for the preparation of financial reports and the use of central financial reporting. Their authority includes annual budget planning in each Ministry in Timor-Leste, based on the State Budget Law No. 2 of 2012 and No. 17 of 2013 (as revised). These laws form the basis for the process of preparing, presenting, and allocating the state budget, which is tailored to government programs (Ho, 2018). Budget allocation in Timor-Leste is program-based, where each budget is set to support the implementation of government programs. The types and sources of data in this study refer to validity and reliability tests. Validity ensures that the questionnaire measures what it is intended to, while reliability ensures consistency of answers over time. Answers are considered reliable if they are consistent each time they are answered.

The calculation of the percentage was conducted by considering the characteristics of the realization component. In general, the measurement of achievement is determined by comparing the realization value with the plnned target and then multiplying the result by 100 percent. When a higher realization value indicates better performance achievement, the achviement level is calculated using the ratio between realization and plan (Dewianawati, 2022).

$$Achievement = \frac{Percentage\ of\ Achievement}{Plan} x\ 100\%$$

$$Achievement = \frac{Target}{Reality} x\ 100\%$$

Source: This formula serves as a descriptive performance indicator to contextualize budget realization. It is treated as an external metric to complement the latent variables analyzed within the PLS-SEM structural model

Measurement of Operational Variables in research using the Likert Scale, according to (Allen & Seaman, 2007), is a method for measuring individual perceptions, attitudes, or opinions towards a statement. This scale consists of 5 points, ranging from strongly agree to strongly disagree. This scale is useful for obtaining ordinal data in social research and helps measure preference or attitude intensity. Based on the source, data collection in this study uses primary data. According to (Sugiyono, 2018), primary data is a data source that directly provides information to data collectors. In this study, primary data was obtained through questionnaires distributed to respondents. Data collection in this study was carried out through interviews, questionnaires, observations, or a combination of the three. The questionnaire contains written questions or statements that are answered by respondents. Questions in the questionnaire can be open or closed statements.

Descriptive analysis in this study uses a Likert scale with a rating of 1 to 5 to measure respondents' answers. The answer rating in the questionnaire consists of the following categories: 1 (Strongly Disagree / STS), 2 (Disagree / TS), 3 (Neutral / N), 4 (Agree / S), and 5 (Strongly Agree / SS). To calculate the index value, a formula is used that involves the frequency of answers to each category, namely F1 for respondents who answered 1, F2 for those who answered 2, F3 for those who answered 3, F4 for those who answered 4, and F5 for those who answered 5. The index value is calculated by summing the value of each category, multiplying by the frequency of answers, then dividing by the total number of respondents (Ferdinand, 2014).

Measurement model (Outer Model): Evaluated via convergent Validity (loading factor > 0.5, ideal >0.7) and Discriminant Validity (Cross-loading and root AVE > 0.5). Reliability is assess using Composite Reliability (>0.6) and Cronbach's Alpha (> 0.761). Structural Model (Inner Model): Hypothesis testing is performed using Bootstrap Resampling to assess path coefficients and the t-test. Based on standard statistical criteria, a relationship is considered significant if the t-statistic is > 1.96 for two-tailed test or > 1.64 for one-tailed, with a corresponding

p-values <0.05. these two methods ensure that the measuring instruments used are consistent and accurate in measuring research constructs (Abdillah & HM, 2015).

Model Structural (Inner Model) Hypothesis testing in PLS is carried out to measure the probability of data using path coefficients. The hypothesis is considered supported if the coefficient of relationship between variables is in accordance with the initial hypothesis and the t-statistic is greater that 1.64 (two-tailed) or 1.96 (one-tailed), which indicates significant results. In additional, a p-value smaller than 0.01, 0.05 or 0.10 also indicates significance. If the p-value is greater than 0.10, the result is considered insignificant (Abdillah & HM, 2015).

Analysis Technique This study uses the PLS-SEM technique to analyze non-linear data. PLS-SEM does not require strict assumptions such as normally distributed data or freedom from multicollinearity. This technique consists of two models: measurement model (for the relationship between manifest and latent variables) and structural model (for the relationship between latent variables). The advantage of PLS is that it does not require a normal distribution and can be used with a small sample, in accordance with this study (Ghozali, 2018).

Hypothesis testing in this study was conducted by examining the values of the Critical Ratio (CR) and Probability (P). A hypothesis was considered accepted when it met the criteria of $P < 0.05$ and $CR > 1.96$. Conversely, a hypothesis was rejected when $P > 0.05$ or $CR < 1.96$. The study employed the Bootstrap Resampling method to obtain more robust parameter estimates, while the t-test was used to assess the magnitude and significance of the effect of each variable. The hypotheses examined in this research included the effect of accountability in the planning stage of Timor-Leste's state budget management on the performance quality of state institutions, the effect of accountability in monitoring state budget management on institutional performance quality, and the effect of accountability in evaluating state budget management on the performance quality of state institutions. Furthermore, the study investigated the moderating role of the Dalan Ba Futuru system in the relationship between accountability in planning, monitoring, and evaluating Timor-Leste's state budget management and the performance quality of state institutions (Sahir, 2021).

RESULT AND DISCUSSION

Data analysis in this study was conducted using the Partial Least Squares Structural Equation Modeling (PLS-SEM) approach with the assistance of SmartPLS software. The PLS-SEM technique was selected because of its ability to accommodate relatively small sample sizes, handle complex research models involving multiple latent constructs, and maximize the explained variance of endogenous variables. In addition, PLS-SEM is particularly suitable for predictive and exploratory research, where the primary objective is to explain and predict relationships among constructs rather than to confirm an established theory (Ghozali, 2018; Hair Jr et al., 2014; Lowry & Gaskin, 2014).

The analytical procedure was performed through two interrelated stages, namely the assessment of the measurement model (outer model) and the structural model (inner model). The measurement model evaluation aims to determine the adequacy of the indicators in measuring their respective latent constructs, while the structural model evaluation examines the relationships among latent variables and the predictive capability of the proposed model.

The validity assessment began with an examination of convergent validity, which reflects the extent to which the indicators of a construct share a high proportion of common variance and effectively represent the underlying latent variable. Convergent validity was evaluated using outer loading values and the Average Variance Extracted (AVE). Outer loadings indicate the strength of the relationship between each indicator and its corresponding construct, with higher values demonstrating stronger indicator reliability. Meanwhile, the AVE value measures the amount of variance captured by a construct relative to the variance attributable to measurement error. According to established guidelines, an AVE value greater than 0.50 indicates that a construct explains more than half of the variance of its indicators, thereby confirming satisfactory convergent validity.

The results of the analysis revealed that all constructs achieved AVE values exceeding the recommended threshold of 0.50. These findings indicate that the indicators employed in this study adequately represent their respective latent variables and possess sufficient convergent validity. Consequently, the measurement model can be considered valid and appropriate for subsequent structural model evaluation and hypothesis testing (Anderson & Gerbing, 1988; Bagozzi & Yi, 2011; Ghozali, 2018).

Table 1. Validity testing results

Variable	Average Variance Extracted (AVE)
Planing (X1)	0.866
Monitoring (X2)	0.842

Evaluation (X3)	0.837
DBFTL (Z)	0.823
Quality of State Institution (Y)	0.848

The results presented in the table indicate that all research constructs have achieved Average Variance Extracted (AVE) values exceeding the recommended threshold of 0.50. This finding demonstrates that each latent construct is capable of explaining more than 50% of the variance of its respective indicators, while the remaining variance is attributable to measurement error. Consequently, the indicators used in this study adequately represent the underlying constructs, confirming that the requirements for convergent validity have been satisfactorily fulfilled. These results provide empirical evidence that the measurement items are sufficiently correlated with the constructs they are intended to measure, thereby supporting the adequacy of the measurement model.

Following the confirmation of convergent validity, the evaluation of construct validity was extended through the assessment of discriminant validity. Discriminant validity is essential for determining whether a construct is empirically distinct from other constructs within the research model. In other words, this assessment ensures that each latent variable captures a unique conceptual domain and does not excessively overlap with other variables. Establishing discriminant validity is particularly important in structural equation modeling, as it enhances the precision of construct measurement and reduces the risk of multicollinearity among latent variables.

In this study, discriminant validity was assessed using the Fornell–Larcker Criterion and Cross Loading analysis. The Fornell–Larcker Criterion evaluates the extent to which a construct shares more variance with its own indicators than with other constructs in the model. This is achieved by comparing the square root of the Average Variance Extracted (√AVE) for each construct with the correlations between that construct and all other constructs. A construct is considered to possess adequate discriminant validity when its √AVE value is greater than its correlations with any other construct. Such a result indicates that the construct is more strongly associated with its own indicators than with indicators belonging to different constructs, thereby confirming its distinctiveness within the model.

The outcomes of the Fornell–Larcker analysis provide important evidence regarding the degree of separation among the constructs included in the study. The detailed results of this assessment are presented in the following table.

Table 2. Fornell-larcker criterion value table

	DBFTL (Z)	AEVA (X3)	Quality of State Institution (Y)	APEM (X2)	APER (X1)
DBFTL	0.907				
Evaluation	0.620	0.915			
Quality of State Institution	0.759	0.594	0.921		
Monitoring	0.608	0.734	0.659	0.917	
Planing	0.451	0.522	0.443	0.654	0.930

Source: Data processing results in the SmartPLS application

The results presented in the table demonstrate that the square root of the Average Variance Extracted (√AVE) for each construct exceeds its correlations with all other constructs in the model. Specifically, the Dalan Ba Futuru (DBFTL) construct achieved a √AVE value of 0.907, which is higher than its correlation coefficients with the remaining variables. Similarly, the Evaluation construct recorded a √AVE value of 0.915, the Institutional Performance Quality (KIN Quality) construct attained a value of 0.921, the Monitoring construct reached 0.917, and the Planning construct obtained the highest √AVE value of 0.930. In each case, the √AVE value surpasses the corresponding inter-construct correlation values.

These findings satisfy the Fornell–Larcker Criterion, indicating that each construct shares a greater proportion of variance with its own indicators than with other constructs included in the research model. Such results provide strong empirical evidence that the latent variables are conceptually and statistically distinct from one another, thereby confirming the uniqueness of each construct. Furthermore, the ability of each construct to demonstrate stronger relationships with its own measurement indicators than with other latent variables suggests that the measurement model possesses adequate discriminant validity.

The fulfillment of the Fornell–Larcker Criterion strengthens the overall quality of the measurement model by reducing concerns regarding construct overlap and multicollinearity among latent variables. Consequently, it can be concluded that the constructs employed in this study exhibit satisfactory discriminant validity, confirming that each variable captures a distinct theoretical dimension and can be reliably used for subsequent structural model assessment and hypothesis testing.

The results of the validity assessment demonstrate that all constructs included in this study have satisfied the requirements of both convergent validity and discriminant validity, confirming that the measurement indicators accurately represent their respective latent variables and are empirically distinct from one another. Following the establishment of validity, the measurement model was further evaluated through reliability testing to assess the internal consistency and stability of the indicators used to measure each construct.

Reliability was examined using two widely accepted measures in Partial Least Squares Structural Equation Modeling (PLS-SEM), namely Composite Reliability (CR) and Cronbach's Alpha. These indicators are employed to determine the extent to which the measurement items consistently capture the same underlying construct. Cronbach's Alpha provides an estimate of internal consistency based on the assumption that all indicators contribute equally to the construct, whereas Composite Reliability offers a more comprehensive assessment by taking into account the actual outer loading values of each indicator. As a result, Composite Reliability is generally considered a more appropriate and robust measure of reliability in PLS-SEM analysis.

The analysis results indicate that all constructs achieved Composite Reliability and Cronbach's Alpha values exceeding the minimum recommended threshold of 0.60. These findings confirm that each construct possesses an acceptable level of internal consistency and reliability. Furthermore, the Composite Reliability values were found to be higher than the corresponding Cronbach's Alpha values across all constructs. This pattern is consistent with the characteristics of PLS-SEM and suggests that the indicators contribute effectively to the measurement of their respective latent variables. The higher Composite Reliability values also indicate that the constructs are measured with a high degree of precision and consistency.

The fulfillment of both reliability criteria provides strong evidence that the measurement instruments used in this study are dependable and capable of producing stable and consistent results. Consequently, the constructs included in the research model can be considered reliable for subsequent structural model evaluation and hypothesis testing. Overall, the findings support the conclusion that the data collection instruments employed in this study possess adequate psychometric properties, ensuring the credibility and robustness of the empirical analysis (Abdillah & HM, 2015; Singh; & Smith, 2006). The detailed values of Composite Reliability and Cronbach's Alpha for each construct are presented in the following table.

Table 3. Composite reliability management value table

	Cronbach's Alpha	Composite Reliability
DBFTL	0.991	0.991
Evaluation	0.988	0.989
Quality of State Institution	0.992	0.993
Monitoring	0.991	0.992
Planing	0.991	0.991

Source: Data processing results in the SmartPLS application

Multicollinearity testing was conducted to assess the extent of correlation among the independent variables included in the research model. This evaluation is essential in structural equation modeling because excessive correlations among predictor variables may lead to estimation bias, unstable parameter coefficients, inflated standard errors, and difficulties in accurately interpreting the individual effects of each independent variable. Therefore, ensuring the absence of multicollinearity is a critical prerequisite for obtaining reliable and valid structural model results.

In this study, multicollinearity was examined using the Variance Inflation Factor (VIF), which measures the degree to which the variance of a regression coefficient is increased due to correlations with other predictor variables. The VIF value serves as an indicator of whether an independent variable shares excessive variance with other variables in the model. According to established guidelines, a VIF value below 5.00 indicates that multicollinearity is not a serious concern, whereas values exceeding this threshold may suggest the presence of problematic collinearity that could compromise the robustness of the model estimates (Ghozali, 2018; Kyriazos & Poga, 2023; Mahmood, 2024).

The results of the analysis reveal that all independent variables exhibit VIF values below the recommended threshold of 5.00. These findings indicate that the predictor variables are not excessively correlated with one another and that each variable contributes unique explanatory information to the model. Consequently, the structural relationships among the constructs can be estimated without significant distortion arising from multicollinearity.

The fulfillment of the VIF criterion confirms that the research model satisfies the multicollinearity assumption and demonstrates an adequate level of independence among the explanatory variables. This result enhances the credibility of the subsequent structural model evaluation and hypothesis testing, as the estimated path coefficients can be interpreted with greater confidence. Therefore, it can be concluded that multicollinearity

does not pose a threat to the validity of the findings, and the proposed model is statistically suitable for further analysis. The detailed VIF values for each independent variable are presented in the following table (Ghozali, 2018).

The results of data processing were utilized to evaluate the proposed hypotheses by examining the path coefficient (β), t-statistic, and P-value obtained from the structural model analysis. Hypothesis testing was conducted to determine the significance and direction of the relationships among the latent variables included in the research model. In accordance with established statistical criteria, a hypothesis was considered supported when the P-value was less than 0.05, indicating that the observed relationship was statistically significant at the 5% significance level. Conversely, hypotheses with P-values greater than 0.05 were regarded as unsupported due to insufficient statistical evidence.

Table 4. Multicollinearity test table

Inner VIF Values	VIF	Description
Planing → Quality of State Institution	1.853	Non Multicollinearity
Monitoring → Quality of State Institution	3.095	Non Multicollinearity
Evaluation → Quality of State Institution	2.429	Non Multicollinearity
DBFTL → Quality of State Institution	1.849	Non Multicollinearity

Source: Data processing results in the SmartPLS application

To ensure robust parameter estimation and significance testing, this study employed the bootstrapping procedure available in SmartPLS. Bootstrapping is a non-parametric resampling technique that repeatedly draws subsamples from the original dataset to generate empirical sampling distributions of the estimated parameters. This approach is particularly advantageous in Partial Least Squares Structural Equation Modeling (PLS-SEM) because it does not require strict assumptions regarding data normality and provides more reliable estimates of standard errors, t-statistics, confidence intervals, and significance levels.

The path coefficient (β) obtained from the analysis reflects both the direction and magnitude of the relationship between constructs. A positive coefficient indicates that an increase in the independent variable is associated with an increase in the dependent variable, whereas a negative coefficient suggests an inverse relationship. The significance of each path coefficient is subsequently evaluated through the bootstrapping results by considering the corresponding t-statistic and P-value. Significant path coefficients provide empirical support for the hypothesized relationships proposed in the conceptual framework.

The outcomes of the bootstrapping analysis offer important evidence regarding the strength and significance of the direct and moderating effects examined in this study. These findings serve as the basis for accepting or rejecting each research hypothesis and contribute to understanding the relationships among the constructs included in the model. The detailed results of the hypothesis testing through the Path Coefficients analysis are presented in the following table.

Table 5. Hypothesis path coefficients value table

	Original Sample (O)	Sample Mean (M)	Standard Deviation (STDEV)	T Statistics (|O/STDEV|)
(X1) Planing →Quality of State Institution	-0.055	-0.042	0.068	Positive and Not Significant.
(X2) Monitoring →Quality of State Institution	0.346	0.316	0.119	Positive and Significant.
(X3) Evaluation →Quality of State Institution	0.025	0.039	0.083	Positive and Not Significant.
(Z) DBFTL →Quality of State Institution	0.582	0.589	0.083	The variable strengthens the relationship between the dependent and independent variables

Source: Data processing results in the SmartPLS application

The measurement model was rigorously assessed to establish its reliability, validity, and overall adequacy for subsequent structural analysis. Reliability evaluation indicated that both Cronbach's Alpha and Composite Reliability values exceeded 0.98 across all constructs, demonstrating an exceptionally high level of internal consistency among the measurement items. While such results confirm the reliability of the instrument, values that are excessively high may also suggest potential item redundancy, indicating that some indicators may capture highly similar aspects of the same construct. This situation can reduce the efficiency of the measurement model and increase the likelihood of common method bias. Therefore, future studies are encouraged to review and refine the measurement items by eliminating overlapping indicators and incorporating more distinct dimensions of the constructs to enhance measurement precision and construct representativeness.

The validity of the measurement model was further examined through discriminant validity assessment using the Heterotrait–Monotrait Ratio of Correlations (HTMT). The HTMT approach is widely recognized as a more sensitive and stringent criterion for evaluating discriminant validity in Partial Least Squares Structural

Equation Modeling (PLS-SEM). The results revealed that all HTMT values were below the recommended threshold of 0.85, indicating that each construct was empirically distinct from the others and captured a unique conceptual domain. These findings provide strong evidence that the latent variables included in the model are sufficiently differentiated and free from substantial construct overlap.

In addition to reliability and validity assessments, the overall model fit was evaluated using the Standardized Root Mean Square Residual (SRMR), which measures the discrepancy between the observed and model-implied correlation matrices. The obtained SRMR value of 0.076 falls below the commonly accepted threshold of 0.08, indicating a satisfactory level of model fit. This result suggests that the proposed model adequately reproduces the observed data structure and demonstrates an acceptable degree of consistency between the theoretical framework and the empirical data.

Furthermore, the predictive capability of the model was assessed through the Stone–Geisser predictive relevance (Q^2) criterion. The analysis showed that all endogenous constructs achieved Q^2 values greater than zero, confirming that the model possesses predictive relevance. Positive Q^2 values indicate that the exogenous variables are capable of predicting the endogenous constructs more accurately than a model without predictive relationships. Consequently, the model not only exhibits satisfactory reliability, validity, and goodness-of-fit but also demonstrates meaningful predictive power, supporting its suitability for explaining and forecasting the relationships among the constructs under investigation.

The findings of this study reveal that, contrary to the initial expectations, accountability in planning and evaluation does not exert a statistically significant positive influence on institutional performance. This result suggests that the existence of accountability mechanisms in the planning and evaluation stages alone may not be sufficient to directly improve organizational outcomes. One possible explanation is the limited capacity of public sector employees to effectively implement accountability practices, which may stem from inadequate training, insufficient technical competencies, and low familiarity with the principles and procedures of public accountability within the Timor-Leste context. As a result, accountability initiatives at the planning and evaluation stages may remain largely procedural and administrative, without being fully translated into tangible improvements in institutional performance.

In contrast, accountability in monitoring was found to have a significant positive effect on institutional performance, with a path coefficient of $\beta = 0.346$. This finding indicates that monitoring activities play a critical role in enhancing organizational effectiveness by facilitating continuous oversight, timely detection of deviations, and immediate corrective actions during the implementation of public financial management processes. Compared with planning and evaluation activities, monitoring mechanisms provide more direct and real-time feedback on operational performance, enabling institutions to respond more effectively to emerging challenges and ensure greater compliance with financial management objectives. Consequently, monitoring accountability appears to be a more influential determinant of short-term institutional performance outcomes.

A particularly important finding of this study is the significant moderating effect of the Dalan Ba Futuru (DBFTL) system on the relationship between accountability practices and institutional performance. The interaction effect between accountability and the DBFTL system was found to be positive and statistically significant ($\beta = 0.582$, $p < 0.001$), indicating that the presence of the system substantially strengthens the impact of accountability mechanisms on organizational performance. This result suggests that the effectiveness of accountability practices is enhanced when supported by an integrated digital platform capable of improving transparency, accessibility, data accuracy, and the timeliness of information related to public financial management.

From a theoretical perspective, these findings provide empirical support for Principal–Agent Theory. The significant moderating role of the DBFTL system demonstrates its capacity to reduce information asymmetry between government institutions, acting as agents, and citizens or other public stakeholders, acting as principals. By facilitating transparent disclosure of financial management processes and improving access to reliable information, the system enhances accountability, strengthens public oversight, and promotes greater trust in governmental institutions. Consequently, the DBFTL system serves not only as a technological innovation but also as an institutional mechanism that improves governance quality and reinforces the effectiveness of accountability practices in achieving higher levels of institutional performance.

CONCLUSION

This study the influence of accountabilty in planning, monitoring and evaluation on the quality of state institution performance in Timor-Leste, while specifically analyzing the moderating role of the Dalan Ba Futuru System (DBFTL). The empirical result demonstrates that while accountability in monitoring significantly influences institutional performance ($\beta = 0.346$, $p = 0.004$), accountability in planning and evaluation does not show a

statistically significant effect in the current model. Crucially, the analysis confirms that the DBFTL system acts as a significant moderating, strengthening the relationship between financial accountability practices and institutional performance (β = 0.582, p < 0.001). These findings contribuite to the literature by extending Principal-Agent Theory and information System Success concepts to public financial management context of Timor-Leste, providing evidence that digital integration essential for reducing information asymmetry. Practically, the study suggests that the government and ANAPMA should prioritize the continue development and full integration of the DBFTL system to enhance budget transparency and effectiveness across all line ministries. The study acknowledges several data, and the cross-sectional design, which limits the ability to infer long-term causal relationships. Future research should seek to expand the institutional scope and employ longitudinal data to further validate these findings and enhance the robustness of conclusions regarding digital governments in emerging public sectors.

REFERENCES

Abdillah, W., & HM, J. (2015). *Partial least square (PLS): Alternatif structural equation modeling (SEM) dalam penelitian bisnis* (1st ed.). Andi Offset.

Alkaraan, F. (2018). Public financial management reform: an ongoing journey towards good governance. *Journal of Financial Reporting and Accounting*, *16*(4), 585–609. https://doi.org/10.1108/JFRA-08-2017-0075

Allen, I. E., & Seaman, C. A. (2007). Likert scales and data analyses. *Quality Progress*, *40*(7), 64–65. https://www.proquest.com/docview/214764202?sourcetype=Magazines

Anderson, J. C., & Gerbing, D. W. (1988). Structural equation modeling in practice: A review and recommended two-step approach. *Psychological Bulletin*, *103*(3), 411–423. https://doi.org/10.1037/0033-2909.103.3.411

Bagozzi, R. P., & Yi, Y. (2011). Specification, evaluation, and interpretation of structural equation models. *Journal of the Academy of Marketing Science*, *40*, 8-34 (2012). https://doi.org/10.1007/s11747-011-0278-x

Birdayanthi, Yusriadi, & Ikmal. (2025). Accountability and transparency in public administration for improved service delivery. *Journal Social Civilecial*, *3*(1), 34–47. https://doi.org/10.71435/610633

Broadbent, J., Dietrich, M., & Laughlin, R. (1996). The development of principal–agent, contracting and accountability relationships in the public sector: Conceptual and cultural problems. *Critical Perspectives on Accounting*, *7*(3), 259–284. https://doi.org/10.1006/cpac.1996.0033

Cuadrado-ballesteros, B., & Bisogno, M. (2022). Budget transparency and financial sustainability. *Journal of Public Budgeting, Accounting & Financial Management*, *34*(6), 210–234. https://doi.org/10.1108/JPBAFM-02-2022-0025

Dewianawati, D. (2022). Analisa kinerja keuangan dengan pendekatan laporan keuangan pada KPPRI. *Owner: Riset & Jurnal Akuntansi*, *6*(1), 454–470. https://doi.org/10.33395/owner.v6i1.591

Ferdinand, A. T. (2014). *Metode penelitian manajemen: Pedoman penelitian untuk penulisan skripsi tesis dan desrtasi ilmu manajemen* (5th ed.). UNDIP Press.

Fatmawati, S., Fitrady, A., & Widodo, T. (2025). Politically driven cycles in fiscal policy: Evidence from disaggregated budgets in middle-income countries. Economies, 13(6), 151. https://doi.org/10.3390/economies13060151.

Ghozali, I. (2018). *Aplikasi analisis multivariate dengan program IBM SPSS 25* (9th ed.). UNDIP Press.

Hair Jr, J. F., Sarstedt, M., Hopkins, L., & Kuppelwieser, V. G. (2014). Partial least squares structural equation modeling (PLS-SEM): An emerging tool in business research. *European Business Review*, *26*(2), 106–121. https://doi.org/10.1108/EBR-10-2013-0128

Ho, A. T.-K. (2018). From performance budgeting to performance budget management: Theory and practice. *Public Administration Review*, *78*(5), 748–758. https://doi.org/10.1111/puar.12915

Kyriazos, T., & Poga, M. (2023). Dealing with multicollinearity in factor analysis: The problem, detections, and Solutions. *Open Journal of Statistics*, *13*, 404–424. https://doi.org/10.4236/ojs.2023.133020

Lowry, P. B., & Gaskin, J. (2014). Partial least squares (PLS) structural equation modeling (SEM) for building and testing behavioral causal theory: When to choose it and how to use it. *IEEE Transactions on Professional Communication*, *57*(2), 123–146. https://doi.org/10.1109/TPC.2014.2312452

Mahmood, S. H. (2024). Estimating models and evaluating their efficiency under multicollinearity in multiple linear regression: A comparative study. *Zanco Journal of Human Sciences*, *28*(5), 264–277. https://doi.org/10.21271/zjhs.28.5.17

Miranda, L. C., Silva, A. J. de M., Filho, J. F. R., & Silva, L. M. da. (2008). An analysis on the comprehension of governmental accounting information informed by the budget balance sheet. *Brazilian Business Review*, *5*(3), 198–217. https://doi.org/10.15728/bbr.2008.5.3.3

Nyman, C., Nilsson, F., & Rapp, B. (2005). Accountability in local government: A principal-agent perspective. *Journal of Human Resource Costing & Accounting*, *9*(2), 123–137. https://doi.org/10.1108/14013380510645397

Pramudiati, N., & Siregar, B. (2018). Local government characteristics, adoption of accrual accounting, and government performance. *Journal of Indonesian Economy and Business*, *33*(2), 153–160.

Sahir, S. H. (2021). *Metodologi penelitian* (4th ed.). Penerbit Kbm Indonesia.

Singh;, P. J., & Smith, A. (2006). An empirically validated quality management measurement instrument. *Benchmarking: An International Journal*, *13*(4), 493–522. https://doi.org/10.1108/14635770610676317

Soares, M. de C., & Wiagustini, N. L. P. (2019). Analisis efektifitas anggaran belanja pada Direção Nacional de Aprovisionamento, Logística e Materiais di Ministério do Comércio, Indústria e Ambiente Timor Leste. *Buletin Studi Ekonomi*, *23*(2), 200–213. https://doi.org/10.24843/BSE.2018.v23.i02.p04

Sugiyono. (2018). *Metode Penelitian Kuantitatif, Kualitatif dan R&D* (23rd ed.). CV. Alfabeta.

Timor-Leste Ministry of Finance. (2024). *Construindo pontes para amanhã: Investir no setor productivo e no capital social*. Dbft.Mof.Tl. https://dbft.mof.gov.tl/pubinfo/quarter_report?lang=tl