International Journal of Science and Engineering Science Research
Volume 2 | Issue 1 | January-March 2026 | ISSN 3049-2793
Peer Reviewed (Referred Journal)
**DOI:** https://doi.org/10.63671/ijsesr.v2i1.69

# Unlocking Predictive Excellence: Key Drivers of Forecasting Accuracy in Knowledge-Based Banking Transformation

**Richard Berimah Twum[123], Abdul-Aziz Ibn Musah[2], Ernestina Hope Turkson[2]**

[1]National Investment Bank PLC, Accra, Ghana

[2]TALI Graduate School, Dominion University College (now Southshore University College), Accra, Ghana

[3]Department of Statistics, University of Cape Coast, Cape Coast, Ghana

Email: richardberimah@gmail.com

## Abstract

In an era where predictive analytics shape strategic banking decisions, understanding what drives forecasting accuracy is no longer optional it is foundational. This study identifies and quantifies the critical factors influencing predictive accuracy within Knowledge-Based Transformation Models (KBTMs), providing actionable intelligence for organizational model development and decision-making.

Drawing on survey data from 310 banking professionals across Ghana's commercial sector, our factor-analytic investigation reveals that predictive accuracy is systematically influenced by four primary drivers: (1) construct-level measurement quality (factor loadings >0.70, AVE >0.60), (2) data structural properties (multicollinearity VIF <3.3, sample adequacy KMO >0.80), (3) model specification complexity (path density, latent variable interactions), and (4) cross-validation strategy (fold configuration, benchmark selection). Knowledge-intensive constructs (e.g., Knowledge Creation, $Q^2$ = 0.834) demonstrate higher sensitivity to measurement quality, while performance outcomes (e.g., Employee Performance, $Q^2$ = 0.191) are more affected by sample characteristics and validation approach.

We establish that predictive tool selection (PLS-Predict vs. CVPAT) should be contingent upon these underlying factors: PLS-Predict excels when measurement quality is high and multicollinearity is present; CVPAT offers advantages when sample constraints or overfitting risks dominate. By translating statistical determinants into strategic guidance, this study empowers banking executives to optimize predictive model development, allocate analytical resources more effectively, and enhance the reliability of knowledge-driven decision-making.

## Keywords

Predictive Analytics, Knowledge-Based Transformation Models (KBTM), Banking Sector Forecasting, Model Validation Techniques

## 1. Introduction

The strategic value of knowledge management in banking depends fundamentally on the ability to anticipate outcomes weather forecasting the impact of knowledge sharing on employee performance or predicting the return on investment in knowledge codification systems (Davenport, 2019; Hair et al., 2017). Yet predictive accuracy is not uniform; it varies systematically based on underlying model characteristics, data properties, and methodological choices (Shmueli et al., 2019; Sharma et al., 2023).

While prior research has established the comparative performance of predictive validation methods (e.g., PLS-Predict vs. CVPAT), less attention has been paid to identifying the factors that explain why predictive accuracy differs across constructs, contexts, and modeling decisions (Liengaard et al., 2021). Understanding these determinants is critical for practitioners seeking to develop robust predictive models and for researchers aiming to advance methodological theory in predictive analytics.

This study addresses Research Objective 5 by investigating the factors that influence predictive accuracy in Knowledge-Based Transformation Models (KBTMs) and examining the implications of using PLS-Predict versus CVPAT for organizational decision-making and model development. Using data from 310

Volume 2 | Issue 1 | January-March 2026 | ISSN 3049-2793
Peer Reviewed (Referred Journal)
**DOI:** https://doi.org/10.63671/ijsesr.v2i1.69

employees across ten commercial banks in Accra, Ghana, we conduct a comprehensive factor-analytic investigation to identify the key drivers of predictive performance.

The contribution is threefold. First, we provide empirical evidence on the relative importance of measurement quality, data structure, model complexity, and validation strategy in determining predictive accuracy. Second, we develop a contingency framework that aligns predictive tool selection with underlying data and model characteristics. Third, we translate statistical findings into practical guidelines for banking practitioners seeking to enhance the reliability of knowledge-driven forecasting.

## 2. Literature Review and Hypotheses Development

### 2.1. Theoretical Foundations of Predictive Accuracy

Predictive accuracy in structural equation modeling is influenced by multiple interrelated factors (Shmueli, 2010; Hair et al., 2019). Drawing on Predictive Modeling Theory and PLS-SEM literature, we identify four primary domains of influence:

Measurement Quality: The reliability and validity of latent variable indicators fundamentally shape predictive performance. Constructs with high factor loadings (>0.70), strong composite reliability (>0.80), and adequate average variance extracted (>0.50) provide more stable estimates for prediction (Hair et al., 2022). Poor measurement quality introduces error variance that attenuates predictive relationships.

Data Structural Properties: Sample size, multicollinearity, and distributional characteristics affect model estimation and prediction. Small samples increase estimation uncertainty; high multicollinearity inflates standard errors; non-normality can bias parameter estimates (Kline, 2016; Sarstedt et al., 2019). PLS-SEM is robust to many of these issues but not immune.

Model Specification Complexity: The number of latent variables, path density, and presence of interaction or higher-order terms influence predictive performance. Overly complex models risk

Volume 2 | Issue 1 | January-March 2026 | ISSN 3049-2793
Peer Reviewed (Referred Journal)
**DOI: https://doi.org/10.63671/ijsesr.v2i1.69**

overfitting; overly simple models may omit critical predictors (Sharma et al., 2023). The optimal complexity depends on sample size and theoretical justification.

Validation Strategy: Cross-validation approach (k-fold configuration, holdout proportion), benchmark selection (Indicator Average vs. Linear Model), and performance metrics ($Q^2$, RMSE, MAE) shape the assessment of predictive accuracy (Liengaard et al., 2021). Different strategies may yield different conclusions about model quality.

### 2.2. Construct-Level Heterogeneity in Predictive Drivers

Not all constructs are equally sensitive to the same predictive drivers. Knowledge-process constructs (e.g., creation, codification) are typically well-specified theoretically and measured with high reliability, making them more sensitive to model specification and validation choices. In contrast, performance outcomes (e.g., employee performance, job satisfaction) are influenced by external, unmeasured factors, making them more sensitive to sample characteristics and data quality (Mbilla et al., 2020).

**H1:** Measurement quality (factor loadings, AVE) will be the strongest predictor of predictive accuracy for knowledge-process constructs.

**H2:** Data structural properties (sample adequacy, multicollinearity) will be the strongest predictor of predictive accuracy for performance-outcome constructs.

**H3:** Model specification complexity will moderate the relationship between validation strategy and predictive accuracy.

**H4:** A contingency framework aligning predictive tool selection with underlying determinants will improve decision-making utility for banking practitioners.

## 3. Methodology

### 3.1. Research Design and Sample

International Journal of Science and Engineering Science Research
Volume 2 | Issue 1 | January-March 2026 | ISSN 3049-2793
Peer Reviewed (Referred Journal)
**DOI:** https://doi.org/10.63671/ijsesr.v2i1.69

This study employed a quantitative cross-sectional survey design. The target population comprised employees from commercial banks operating in the Accra Metropolitan District, Ghana.

**Sample Size:** 310 valid responses obtained from employees across ten commercial banks.

Response Rate: 98.41%.

**Demographics:** 58% male, 42% female; majority (56.1%) with 5-10 years of work experience.

**Table 4.1:** Demographic Characteristics of Bank Employees

| Demographics | Frequency | Percent |
|---|---|---|
| **Job Rank** | | |
| Junior | 83 | 26.8 |
| Middle Management | 54 | 17.4 |
| Officer | 162 | 52.3 |
| Senior Management | 11 | 3.5 |
| **Total** | **310** | **100.0** |

## 3.2. Measurement Instruments

Data were collected using a structured questionnaire adapted from validated scales:

**Knowledge Management Constructs:** Six dimensions (Creation, Acquisition, Sharing, Application, Codification, Retention) adapted from Kianto (2008) and Nonaka & Takeuchi (1995).

**Employee Performance (EP):** Assessed using items focusing on task effectiveness and productivity, adapted from Koopmans et al. (2019).

**Job Satisfaction (DJS):** Measured using items from Lee, Kim, & Park (2020).

**Government Policies (GP):** Adapted from Garcia & Martinez (2023).

Volume 2 | Issue 1 | January-March 2026 | ISSN 3049-2793
Peer Reviewed (Referred Journal)
**DOI: https://doi.org/10.63671/ijsesr.v2i1.69**

### 3.3. Analytical Procedure

**Factor-Analytic Investigation:** Principal component analysis with varimax rotation to identify underlying factors influencing predictive accuracy.

**Predictor Variables:** Measurement quality indicators (loadings, AVE, CR), data properties (KMO, VIF, sample size), model complexity (path count, interaction terms), validation strategy (fold number, benchmark type).

**Outcome Variable:** Predictive accuracy ($Q^2$ Predict, RMSE) for each latent construct.

**Contingency Analysis:** Moderated regression to test H3; decision-tree analysis to develop practical guidelines for H4.

## 4. Results

### 4.1. Factor-Analytic Identification of Predictive Drivers

Principal component analysis extracted four factors explaining 78.3% of variance in predictive accuracy metrics:

| Factor | Key Indicators | Eigenvalue | % Variance |
|---|---|---|---|
| **F1: Measurement Quality** | Factor loadings, AVE, CR | 4.21 | 32.4% |
| **F2: Data Structure** | KMO, VIF, sample adequacy | 2.87 | 22.1% |
| **F3: Model Complexity** | Path density, interaction terms | 1.94 | 14.9% |
| **F4: Validation Strategy** | Fold configuration, benchmark | 1.15 | 8.9% |

Measurement Quality emerged as the strongest predictor for knowledge-process constructs ($\beta = 0.67$, $p < 0.001$), supporting H1.

International Journal of Science and Engineering Science Research
Volume 2 | Issue 1 | January-March 2026 | ISSN 3049-2793
Peer Reviewed (Referred Journal)
**DOI:** https://doi.org/10.63671/ijsesr.v2i1.69

Data Structure was the dominant predictor for performance-outcome constructs (β = 0.54, p = 0.003), supporting H2.

Model Complexity significantly moderated the validation strategy-accuracy relationship ($\Delta R^2 = 0.11$, $p = 0.018$), supporting H3.

## 4.2. Construct-Level Sensitivity Analysis

Regression analysis revealed differential sensitivity across construct types:

| Construct Type | Most Sensitive To | Least Sensitive To |
|---|---|---|
| **Knowledge Processes (KC, KCO, KR)** | Measurement Quality (β = 0.67) | Validation Strategy (β = 0.12) |
| **Performance Outcomes (EP, DJS)** | Data Structure (β = 0.54) | Model Complexity (β = 0.19) |
| **Contextual Factors (GP)** | Validation Strategy (β = 0.41) | Measurement Quality (β = 0.28) |

Knowledge Creation (KC) showed the highest sensitivity to measurement quality: a 0.10 increase in average factor loading corresponded to a 0.15 increase in $Q^2$ Predict ($p < 0.001$). Employee Performance (EP) showed the highest sensitivity to sample adequacy: KMO > 0.85 was associated with 23% lower RMSE ($p = 0.007$).

## 4.3. Contingency Framework for Tool Selection

Decision-tree analysis identified optimal conditions for PLS-Predict vs. CVPAT:

| Condition | Recommended Method | Expected Accuracy Gain |
|---|---|---|

| | | |
|---|---|---|
| High measurement quality (AVE > 0.65) + Multicollinearity (VIF > 3.0) | PLS-Predict | +18% $Q^2$, -22% RMSE |
| Low sample adequacy (KMO < 0.75) + Small sample (n < 200) | CVPAT | +12% robustness, -15% overfitting risk |
| High model complexity (>15 paths) + Interaction terms | PLS-Predict | +14% predictive relevance |
| Performance-focused outcomes + External validity priority | CVPAT | +19% generalizability |

This framework supports H4, demonstrating that aligning method selection with underlying determinants improves practical utility.

### 4.4. Hypothesis Testing Summary

**H1 Supported:** Measurement quality was the strongest predictor of predictive accuracy for knowledge-process constructs.

**H2 Supported:** Data structural properties dominated for performance-outcome constructs.

**H3 Supported:** Model complexity significantly moderated the validation strategy-accuracy relationship.

**H4 Supported:** A contingency framework improved decision-making utility for banking practitioners.

## 5. Discussion

### 5.1. Interpretation of Predictive Determinants

The findings reveal that predictive accuracy is not monolithic but systematically shaped by four interrelated domains. Measurement quality emerged as paramount for knowledge-process constructs,

International Journal of Science and Engineering Science Research
Volume 2 | Issue 1 | January-March 2026 | ISSN 3049-2793
Peer Reviewed (Referred Journal)
**DOI: https://doi.org/10.63671/ijsesr.v2i1.69**

aligning with theoretical expectations that well-specified, reliably measured latent variables yield more stable predictions (Hair et al., 2022). For performance outcomes, data structure dominated, reflecting the influence of external, unmeasured factors that amplify the importance of sample adequacy and multicollinearity management (Mbilla et al., 2020).

The moderating role of model complexity underscores a critical trade-off: complex models capture nuanced relationships but require larger samples and careful validation to avoid overfitting (Sharma et al., 2023). This finding challenges one-size-fits-all approaches to predictive modeling and supports a contingency perspective.

### 5.2. Theoretical Implications

This study extends predictive modeling theory by identifying construct-level heterogeneity in the drivers of predictive accuracy. It suggests that future PLS-SEM research should report determinant profiles alongside predictive metrics, enabling more nuanced theoretical development. Furthermore, it highlights the value of factor-analytic approaches in diagnostic model assessment.

### 5.3. Practical Implications

For Banking Leaders: Prioritize measurement quality when forecasting knowledge-process outcomes; invest in data quality and sample adequacy when predicting performance outcomes. Use the contingency framework to select validation methods aligned with model characteristics.

For Analytics Teams: Conduct determinant profiling during model development: assess measurement quality, data structure, and complexity before selecting validation strategy. Document determinant profiles to support model refinement and stakeholder communication.

For Policymakers: Recognize that predictive reliability depends on underlying data and model properties. Encourage transparency in reporting determinant profiles to support evidence-based policy evaluation.

### 5.4. Strategic Value of Determinant-Aware Modeling

Understanding what drives predictive accuracy transforms model development from a technical exercise into a strategic capability. Banks that systematically assess and optimize predictive determinants can: (1) allocate resources more efficiently (e.g., invest in measurement refinement vs. sample expansion based on construct type); (2) reduce forecasting error in critical decisions (e.g., knowledge management investments); and (3) enhance organizational learning by building more reliable predictive systems.

## 6. Conclusion and Recommendations

### 6.1. Conclusion

This study establishes that predictive accuracy in Knowledge-Based Transformation Models is systematically influenced by four primary determinants: measurement quality, data structure, model complexity, and validation strategy. These factors exhibit construct-level heterogeneity, with knowledge processes most sensitive to measurement quality and performance outcomes most sensitive to data properties. A contingency framework aligning predictive tool selection with these determinants improves practical utility for banking practitioners.

### 6.2. Recommendations

Adopt Determinant Profiling: Integrate assessment of measurement quality, data structure, and model complexity into standard model development protocols.

Implement Contingency-Based Tool Selection: Use the proposed framework to select PLS-Predict or CVPAT based on underlying determinant profiles.

Invest in Construct-Specific Optimization: Prioritize measurement refinement for knowledge-process constructs; focus on data quality and sample adequacy for performance outcomes.

International Journal of Science and Engineering Science Research
Volume 2 | Issue 1 | January-March 2026 | ISSN 3049-2793
Peer Reviewed (Referred Journal)
**DOI:** https://doi.org/10.63671/ijsesr.v2i1.69

Enhance Transparency in Reporting: Report determinant profiles alongside predictive metrics to support model interpretation and replication.

### 6.3. Limitations and Future Research

This study is limited by its cross-sectional design and reliance on self-reported data. Future research should: (1) employ longitudinal designs to assess determinant stability over time; (2) incorporate objective performance metrics to validate self-reports; (3) extend the investigation to other sectors and digital transformation contexts; and (4) explore machine learning approaches to automate determinant profiling and tool selection.

### References

1. Davenport, T. H. (2019). The AI Advantage: How to Put the Artificial Intelligence Revolution to Work. MIT Press.
2. Garcia, A., & Martinez, L. (2023). Government policies and knowledge management in the banking sector. Journal of Policy Research, 48(4), 212-231.
3. Hair, J. F., et al. (2017). PLS-SEM: Indeed, a silver bullet. Journal of Marketing Theory and Practice, 25(2), 139-151.
4. Hair, J. F., et al. (2019). When to use and how to report the results of PLS-SEM. European Business Review, 31(1), 2-24.
5. Hair, J. F., et al. (2022). Partial Least Squares Structural Equation Modeling (PLS-SEM) Using R: A Workbook. Springer.
6. Kianto, A. (2008). Development and validation of a survey instrument for measuring organisational knowledge management. Knowledge and Process Management, 15(1), 1-13.
7. Kline, R. B. (2016). Principles and Practice of Structural Equation Modeling (4th ed.). Guilford Press.

Volume 2 | Issue 1 | January-March 2026 | ISSN 3049-2793
Peer Reviewed (Referred Journal)
**DOI:** https://doi.org/10.63671/ijsesr.v2i1.69

8. Koopmans, L., et al. (2019). Individual work performance questionnaire. Journal of Occupational Medicine.
9. Lee, S., Kim, J., & Park, H. (2020). The impact of knowledge management on job satisfaction in the banking sector. Journal of Knowledge Management, 24(3), 567-584.
10. Liengaard, B. D., et al. (2021). Prediction: Covariance-based or Partial Least Squares Structural Equation Modeling? Journal of Travel Research, 60(7), 1564-1583.
11. Mbilla, N., et al. (2020). Impact of KM practices on bank performance in Ghana. African Journal of Business Management.
12. Nonaka, I., & Takeuchi, H. (1995). The Knowledge-Creating Company. Oxford University Press.
13. Sarstedt, M., et al. (2019). When to use and how to report the results of PLS-SEM. European Business Review.
14. Sharma, P. N., et al. (2023). Predictive model assessment and selection in composite-based modeling using PLS-SEM. European Journal of Marketing, 57(6), 1662-1677.
15. Shmueli, G. (2010). To explain or to predict? Statistical Science, 25(3), 289-310.
16. Shmueli, G., et al. (2019). Predictive model assessment in PLS-SEM. European Journal of Information Systems, 28(1), 1-18.