

Artificial Intelligence-Based Procurement Risk Screening and Single-Bid Competition in Nigeria

Yusuf, A. A¹, Abdulkadir, A. B², Aransiola, O. T³.

^{1&2}Department of Accounting, Faculty of Arts, Management, & Social Sciences
Abdulrasaq Abubakar Toyin University, Nigeria

³Department of Accounting, Faculty of Management Sciences
Al-Hikmah University, Ilorin, Nigeria

*Corresponding Author: yusufabubakarahmed@gmail.com

ABSTRACT

This study assesses whether artificial-intelligence methods add practical value when used to screen single-bid competition risk in Nigerian federal procurement. The evidence comes from Bureau of Public Procurement records published in Open Contracting Data Standard format. The underlying file contains 98,866 compiled contracting processes, of which 9,124 satisfy the study's analytical requirements: a positive tenderer count and an interpretable tender-closing date between 2017 and 2024. A single-bid process is treated as a signal of weak observed competition that may justify closer examination; it is not treated as evidence of corruption or procurement misconduct. To keep the information sequence intact, each buyer-history variable is calculated only from procurement processes that closed before the process being scored. Random forest and logistic regression are estimated with observations from 2017-2022 and evaluated on a later 2023-2024 holdout, together with simpler benchmarks based on buyer history and procurement method. Both multivariable models show useful but moderate discrimination (ROC AUC = 0.761). Random forest records average precision of 0.633 and a Brier score of 0.190, while logistic regression records average precision of 0.619 and a lower Brier score of 0.187. A buyer-cluster bootstrap based on 1,000 replications provides no statistically meaningful evidence that the two AUCs differ, and logistic regression has a calibration slope closer to one. The findings therefore support AI-assisted ranking and prioritization for audit attention, not automated detection of wrongdoing, and they show that greater model complexity does not automatically produce better assurance decisions.

KEYWORDS: Artificial Intelligence; Audit Analytics; Public Procurement; Single Bidding

I. INTRODUCTION

Public procurement is the route through which government expenditure is translated into goods, works and services. It therefore has a direct bearing on financial control, value for money and the accountability of public institutions. Nigeria's Public Procurement Act 2007 places competition and transparency among the central principles of public purchasing, while allowing other procedures in circumstances recognised by the law. The Bureau of Public Procurement also manages the Nigeria Open Contracting Portal, which releases procurement information in the Open Contracting Data Standard. Machine-readable records of this kind can do more than satisfy disclosure requirements. Properly organised, they can help audit units sort large volumes of procurement activity and decide which cases should receive attention first. Interest in artificial intelligence within accounting and auditing has grown partly because the volume of administrative data now exceeds what can reasonably be examined through manual review alone. Nigerian accounting researchers have begun to study AI in areas such as risk management and human-resource accounting (Adedire-Ampitan et al., 2024; Odiase & Uagbale-Ekatah, 2025), but much of the domestic evidence remains perception-based. That leaves a more operational question largely unanswered: when these techniques are applied to

Nigerian administrative procurement records themselves, do they improve the quality of risk screening? The international accounting literature also cautions against treating predictive accuracy as the only test of usefulness. In assurance work, a model must be explainable to its users, open to audit, compared with credible alternatives and supported by dependable information (Kokina et al., 2025; Li & Goel, 2025; Stratopoulos & Wang, 2025).

This study addresses that practical gap with federal procurement records from BPP/NOCOPO. Single bidding is used because the reported number of tenderers provides a direct indication of competition at the process level. It is not treated as a finding of fraud, corruption or illegality. Instead, it is used as an initial screening signal that may assist auditors and oversight bodies in deciding where further inquiry is warranted. The main comparison is intentionally simple: does random forest, despite its greater complexity, improve materially on logistic regression and on basic historical benchmarks when each model works from the same information set? The analysis is organised around three objectives. First, it establishes how frequently single-bid procurement occurs within the cohort for which tender-closing dates can be observed and considers how that incidence changes over time. Second, it examines whether information about a buyer's prior behaviour, when combined with features of the current tender, provides better discrimination than simple historical rules. Third, it compares random forest with logistic regression on a later holdout period to determine whether the added complexity of the former yields a material predictive advantage. These objectives translate into three practical questions: how does single bidding vary across the analytical period; what additional screening information is gained from structured procurement characteristics; and does random forest outperform logistic regression when both are tested out of time? The paper contributes in two related ways. On the methodological side, it shows that claims about AI performance are shaped by choices made before estimation begins. The unit being predicted, the point in time at which information becomes available, the benchmark selected for comparison and the calibration of predicted probabilities all affect the meaning of a model's reported performance. On the empirical side, the study uses a large body of Nigerian administrative procurement records rather than asking respondents to describe their perceptions of AI. For accountants, internal auditors, procurement officials and anti-corruption institutions, this distinction matters because it helps identify when AI-assisted screening produces genuinely useful information and when a simpler model may offer similar performance with greater transparency and easier review.

II. RELATED WORKS

A. Public Procurement, Competition and Forensic Red Flags

Open contracting seeks to make public procurement more visible by publishing structured information from planning through tendering, award, contracting and implementation. Within Nigeria, BPP presents NOCOPO as a mechanism for strengthening disclosure, competition and accountability. Competition is also a core principle of the Public Procurement Act 2007. Against that institutional background, a process receiving only one tenderer is a meaningful screening signal because there is no observed rival bid at tender close against which price, quality or delivery terms can be competitively tested. Comparable procurement research has therefore used single bidding, restricted competition, repeated supplier relationships and unusual procedures to identify cases that may warrant closer scrutiny (Fazekas et al., 2016; Fazekas & Kocsis, 2020; Gallego et al., 2021). The signal is deliberately interpreted narrowly. A single bid can arise for legitimate reasons, including a thin or specialised supplier market, technical specificity, emergency conditions, geography or other lawful procurement circumstances. It is therefore a competition-risk indicator for prioritizing review, not a finding of corruption, collusion, illegality or poor value for money.

B. AI, Audit Analytics and Accounting Information Systems

Accounting analytics increasingly uses machine learning to classify, predict and prioritize risk. Fedyk et al. (2022) show that AI adoption can affect audit processes, while recent accounting-information-systems research emphasizes transparency, explainability, robustness and auditability (Kokina et al., 2025; Li & Goel, 2025). Stratopoulos and Wang (2025) further distinguish accounting centered AI research from studies in which technical novelty becomes the main objective. In a public-sector assurance setting, an algorithm should therefore be judged not only by predictive accuracy, but also by whether its information inputs are legitimate, whether the result can be explained and whether complexity improves decision value over simpler alternatives.

C. Nigerian Evidence and Research Gap

Recent Nigerian accounting studies demonstrate growing interest in AI. Adedire-Ampitan et al. (2024) report evidence on AI and risk management in Nigerian banking, while Odiase and Uagbale-Ekatah (2025) examine AI and human-resource accounting. These studies establish the relevance of AI to Nigerian accounting practice but use primary-data designs. The present study extends this literature by using machine-readable public administrative records and by evaluating predictive models with an out-of-time design. It also addresses a gap in Nigerian public-sector accounting research by linking open contracting, audit analytics and model governance.

D. Agency Theory and Monitoring

Agency theory provides the principal theoretical lens. Public officials and procurement entities act as agents in spending resources ultimately belonging to citizens and the state. Information asymmetry can limit the ability of principals, legislatures, auditors and citizens to monitor procurement decisions (Jensen & Meckling, 1976). Open contracting reduces part of this asymmetry by making process information observable, while analytics can reduce monitoring costs by ranking cases for audit attention. However, an opaque or poorly validated algorithm can create a new information asymmetry between model developers and users. The study therefore treats transparent benchmarking and calibration as extensions of the monitoring function: complex models should be retained only when they provide demonstrable additional value.

III. METHODOLOGY

A. Research Design and Data Source

The study adopts a retrospective longitudinal predictive design. Data come from the Nigeria BPP dataset distributed through the Open Contracting Partnership Data Registry and sourced from NOCOPO. In OCDS terminology, each observation in the registry is a compiled contracting process. The all-time gzip snapshot used for the analysis contains 98,866 compiled processes. The registry itself warns of date inconsistencies and explains that many release dates may reflect internal updates rather than first publication; consequently, this study does not claim a live historical deployment test.

B. Sample Construction and Unit of Analysis

The target variable requires an observed positive number of tenderers. There are 17,623 such processes, representing 17.8% of the 98,866-process raw snapshot. Of these, 9,210 contain an interpretable tender-closing date; restricting tender-closing dates to 2017-2024 yields 9,124 contracting processes across 241 normalised buyers. The final analytical cohort therefore represents 51.8% of processes with a positive tenderer count but 9.2% of all compiled processes in the snapshot. Award value is not used as a sample condition, predictor or outcome. The contracting process not the individual award is the analytical unit because number of tenderers

is defined at tender/process level. Processes containing multiple award objects remain one observation. This filtering has an important implication for external validity. The excluded records are not a random holdout from the same fully observed population: most are excluded because the single-bid outcome cannot be defined from a positive tenderer count or because chronological ordering cannot be established from an interpretable tender-close date. The final cohort may therefore over-represent procuring entities, periods or process types with more complete OCDS reporting. Because the target itself is unavailable for most excluded records, the study cannot test whether their single-bid incidence differs systematically from the retained cohort. The results should consequently be interpreted as conditional on Nigerian federal contracting processes for which competition and timing fields are sufficiently complete for leakage-controlled screening, rather than as prevalence or predictive-performance estimates for all federal procurement activity.

Table 1: Sample Construction

Stage	Processes	Rationale	Retention of raw snapshot
Raw compiled processes	98,866	All OCIDs in downloaded snapshot	100.0%
Positive tenderer count	17,623	Single-bid outcome definable	17.8%
Interpretable tender-closing date	9,210	Chronological ordering possible	9.3%
Tender close in 2017-2024	9,124	Final analytical cohort	9.2%
Normalised buyers	241	Buyer-history construction	—

Source: Authors' analysis of Nigeria BPP/NOCOPO OCDS data.

C. Variables and Information Boundary

Single-bid equals one when number of tenderers equals one and zero otherwise. It is a tender/process-level competition-risk label rather than a misconduct label. The main predictors are the buyer's prior single-bid rate, the logarithm of the buyer's strictly prior process count, procurement method details, plausible tender-window duration and a missing-duration indicator. Historical buyer variables are updated only after all processes sharing the same tender-closing date are scored, preventing same-date leakage. Tender windows outside 1-365 days are not used as continuous values. Procurement method details are retained categorically; categories such as selective tendering, direct procurement, emergency and sole source are descriptive procedure types and are not labelled as misconduct.

Table 2: Variables and Analytical Timing

Variable	Definition / timing
Single bid	1 if number of tenderers = 1; tender/process-level target
Prior buyer single-bid rate	Share of same buyer's processes on strictly earlier tender-close dates that were single bid
Prior buyer process volume	Count of same buyer's processes on strictly earlier tender-close dates; $\log(1+x)$
Procedure category	procurement Method Details grouped/categorical; no misconduct presumption
Tender-window duration	Tender start-to-close days retained only when 1-365 days

Variable	Definition / timing
Duration missing	Indicator for unavailable/implausible tender-window duration

Source: Authors' analysis of Nigeria BPP/NOCOPO OCDS data

D. Model Implementation and Reproducibility

Both classifiers use the same information set and the same chronological training/holdout boundary. The procedure-category field is passed through the archived preprocessing pipeline as a categorical model input, while buyer-history variables and tender-window information are generated before estimation using only information available at the relevant tender-close date. The 2023-2024 terminal holdouts is reserved for final out-of-time evaluation and is not used as an information source for constructing buyer histories. This common information boundary ensures that differences in model performance reflect the estimators rather than access to different predictors. For full reproducibility, the executable analysis package cited in the Data Availability statement is the authoritative specification of implementation details, including the exact random-forest constructor arguments, random seed, categorical transformation, logistic-regression specification and model-fitting sequence. These settings are preserved in code rather than inferred retrospectively from reported performance. The revised manuscript therefore makes the preprocessing and information boundary explicit while directing readers to the archived script for the exact hyper parameter values and parameter-selection instructions used to reproduce Tables 4-6 and Figures 2-3.

E. Statistical and machine-learning procedures

The principal training period is 2017-2022 and the terminal test period is 2023-2024. A random forest classifier is compared with logistic regression using identical predictors and the common preprocessing/information boundary described above. Two simpler benchmarks are also evaluated: prior buyer single-bid rate alone and empirical procurement-method rate. Performance is assessed using ROC area under the curve (AUC), average precision (AP), Brier score, calibration slope and highest risk decile lift. AUC is interpreted as discrimination rather than overall accuracy; AP is included because the positive class is not balanced; and the Brier score and calibration slope assess the quality of predicted probabilities. Because observations are clustered within buyers, uncertainty is assessed with 1,000 paired buyer-cluster bootstrap replications. Rolling-origin tests evaluate temporal stability from 2020 to 2024, and a cold-start analysis separately examines buyers not observed in the training period. Random-forest permutation importance is computed on the terminal holdout. The analysis is predictive rather than causal, and the terminal holdout is used to evaluate not to retrospectively justify model complexity.

IV. RESULTS AND DISCUSSION

This section presents and discusses the empirical findings from the analysis of Nigerian public procurement records. It first examines the descriptive patterns of single-bid procurement across tender-closing years and procurement categories, before presenting the results of the machine-learning models used for risk screening. The findings are interpreted in relation to the identified procurement characteristics, with particular attention to model performance, predictive discrimination, and the practical implications of using AI-assisted approaches to identify procurement processes that may warrant further review.

A. Descriptive Evidence

Table 3 presents the annual distribution of single-bid procurement processes within the analytical cohort from 2017 to 2024. It shows the number of procurement processes recorded in each year alongside the corresponding single-bid rate, providing an overview of how the incidence of single-bid procurement changed over the study period. The variation observed across the years provides useful descriptive evidence of changes in procurement competition and supports the need for temporal analysis when developing models for procurement risk screening.

Table 3: Single-Bid Procurement by Tender-Closing Year

Year	Processes	Single-bid Rate
2017	295	39.0%
2018	796	36.4%
2019	1,227	36.8%
2020	1,462	44.0%
2021	1,534	32.9%
2022	1,832	44.3%
2023	1,266	28.3%
2024	712	30.8%

Source: Authors' analysis of Nigeria BPP/NOCOPO OCDS data.

The final cohort contains 9,124 contracting processes. Single-bid prevalence varies materially across years, indicating that the prediction problem is not temporally stationary. This reinforces the need for out-of-time validation rather than random train-test splits. Figure 1 presents the annual pattern.

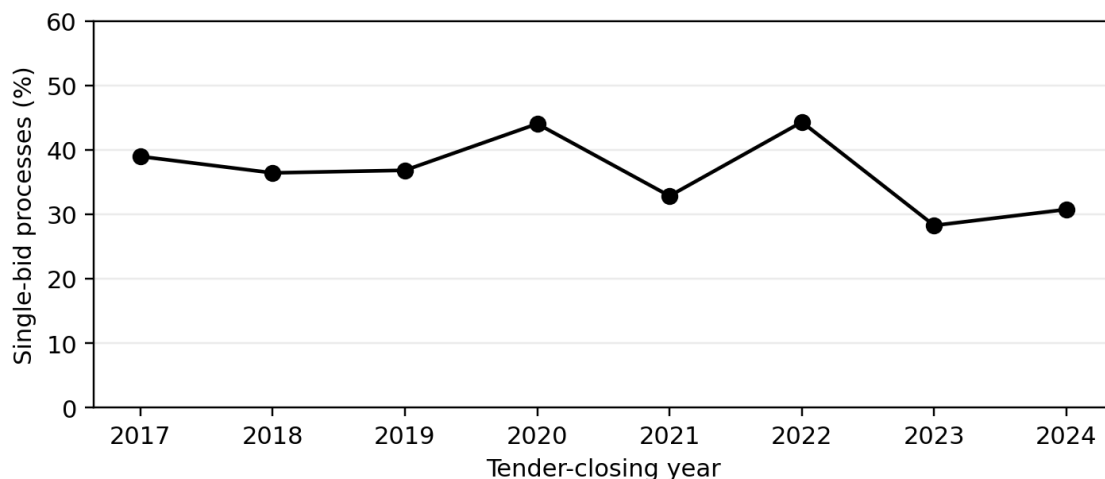


Figure 1: Annual Single-Bid Prevalence in the Analytical Cohort.

Source: Authors' analysis of Nigeria BPP/NOCOPO OCDS data.

B. Terminal Holdout Performance and Model Comparison

As presented in Table 4, Random forest and logistic regression produce almost identical discrimination: AUCs are 0.761 and 0.761 respectively. These values indicate useful but moderate ranking ability rather than highly accurate classification. Both models outperform the buyer-history-only benchmark (AUC 0.705) and the procurement-method benchmark (AUC 0.659), showing that combined structured context adds screening information. However, logistic regression records the lower Brier score (0.187 versus 0.190) and a calibration slope closer to the ideal value of one (0.951 versus 1.488). The Brier score evaluates the squared

error of predicted probabilities, so lower values are preferable. A calibration slope near one indicates that probability differences are appropriately scaled; the random-forest slope above one suggests that its holdout probabilities are less well scaled than those from logistic regression. Thus, the models are more defensible as ranking and prioritisation tools than as stand-alone decision systems, and the more complex model does not provide a clear practical advantage.

Table 4: Model Performance on 2023-2024 Tender Closures

Model	ROC AUC	Average precision	Brier	Calibration slope
Random forest	0.761	0.633	0.190	1.488
Logistic regression	0.761	0.619	0.187	0.951
Buyer-history rate	0.705	0.526	0.204	0.332
Method empirical rate	0.659	0.478	0.200	1.757

Source: Authors' analysis of Nigeria BPP/NOCOPO OCDS data.

As presented in Table 5, the paired buyer-cluster bootstrap confirms that the random-forest-minus-logistic AUC difference is -0.001 with a 95% confidence interval from -0.029 to 0.017. The corresponding AP-difference interval also includes zero. There is therefore no statistical basis for claiming superior predictive performance from random forest in the terminal holdout. The uncertainty intervals also reinforce the need to describe the observed discrimination as moderate and screening-oriented rather than as evidence of a highly accurate detector.

Table 5: Buyer-Cluster Bootstrap Uncertainty (1,000 Replications)

Statistic	Estimate	95% CI
Random-forest AUC	0.761	[0.675, 0.864]
Logistic AUC	0.761	[0.676, 0.870]
Buyer-history AUC	0.705	[0.594, 0.821]
RF - logistic AUC	-0.001	[-0.029, 0.017]
RF - logistic AP	0.013	[-0.043, 0.049]

Source: Authors' analysis of Nigeria BPP/NOCOPO OCDS data.

C. Temporal Stability, Cold-Start Performance and Feature Importance

Rolling validation shows that performance changes across test years and that neither model dominates consistently. This is important for governance: model performance should be monitored over time rather than assumed to remain stable after deployment. The 2023-2024 holdout includes 30 buyers absent from the training period (177 processes). Random forest AUC is 0.821 for this small cold-start group and 0.760 for previously observed buyers; because the new-buyer subgroup is small, the result is treated as exploratory rather than definitive.

Table 6: Rolling-Origin Discrimination

Test year	N	Single-bid prevalence	RF AUC	Logit AUC
2020	1,462	44.0%	0.814	0.814
2021	1,534	32.9%	0.711	0.665
2022	1,832	44.3%	0.720	0.795
2023	1,266	28.3%	0.760	0.759
2024	712	30.8%	0.783	0.743

Source: Authors' analysis of Nigeria BPP/NOCOPO OCDS data.

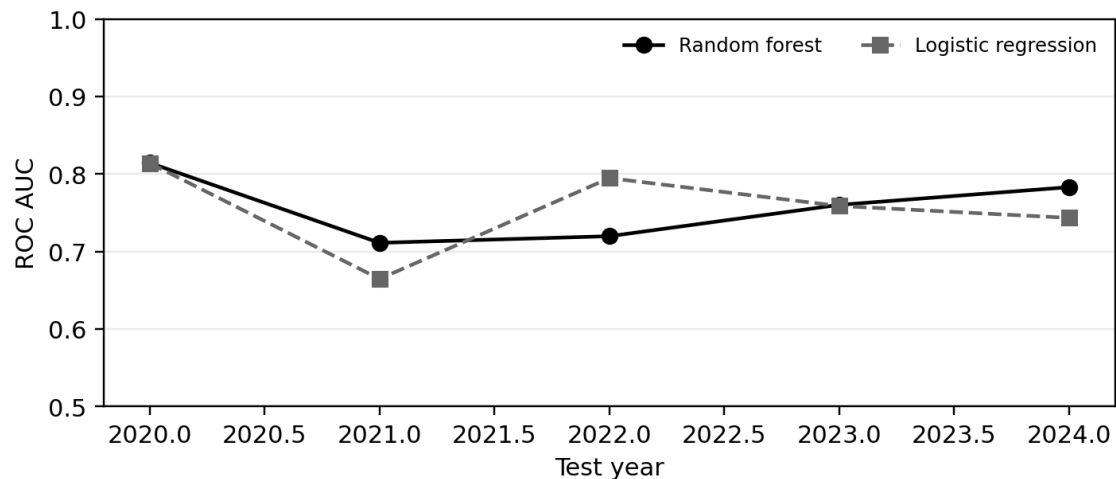


Figure 2: Rolling-Origin Roc Auc for Random Forest and Logistic Regression.

Source: Authors' analysis of Nigeria BPP/NOCOPO OCDS data.

Permutation analysis shows that prior buyer single-bid rate is the most influential random-forest feature, followed by procedure category and prior buyer process volume. Tender-window duration adds comparatively little predictive value in the terminal holdout. Figure 3 therefore reinforces the substantive interpretation that persistent buyer-level competition patterns are the principal source of predictive information.

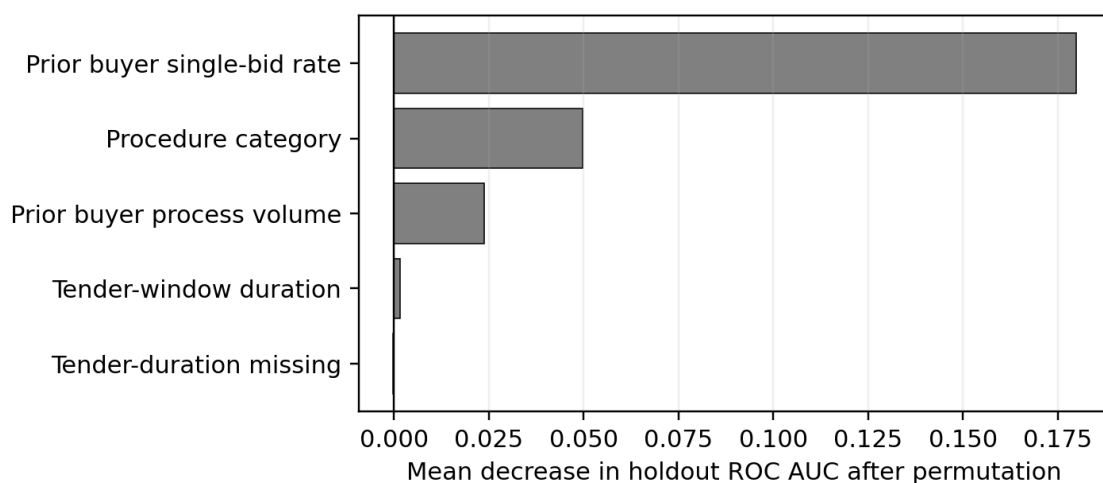


Figure 3: Random forest permutation importance on the terminal holdout.

Source: Authors' analysis of Nigeria BPP/NOCOPO OCDS data.

D. Discussion

The results point to three broad observations. The first is that the procurement records contain information that can assist risk prioritisation. Models that combine earlier buyer behaviour with tender characteristics perform better than benchmarks built from buyer history or procurement method alone. In practical audit work, this means that screening improves when information from different parts of the procurement record is considered together rather than in isolation. The second observation is that technical sophistication should not be confused with decision usefulness. Random forest does not significantly outperform logistic regression when both models use the same predictors and are assessed on the same later-period holdout. The logistic model is also better calibrated: its lower Brier score and calibration slope of 0.951 mean that its probability estimates are more closely aligned with observed outcomes than the random-

forest probabilities. This matters in assurance practice because a risk score may be used to determine how many cases enter an audit queue or where a review threshold is set. Better-calibrated probabilities make those resource-allocation decisions easier to interpret and defend; a nominal probability should correspond reasonably closely to the frequency of the screened outcome among comparable cases. For Nigerian public-sector assurance, the result therefore argues for proportionality in model choice. Where a transparent method discriminates as well as a more elaborate one and is better calibrated, interpretability, ease of scrutiny and lower implementation cost become substantive advantages rather than secondary conveniences. This position is consistent with accounting research that places explainability, reliability and human oversight at the centre of AI-enabled audit work (Kokina et al., 2025; Li & Goel, 2025).

The third observation concerns buyer history. Prior single-bid rate carries the greatest random-forest importance, indicating that procurement competition has some persistence across processes. Agency theory offers one way of understanding this pattern: repeated, observable outcomes may reduce monitoring costs by helping oversight bodies direct scarce audit effort toward buyers with a record of weaker competition. The statistical model does not reveal why that persistence occurs. Supplier-market structure, specialised procurement, geography, emergency conditions, reporting practice and organisational capability are all plausible explanations. The resulting score should therefore be read as a competition-risk signal and not as a diagnosis of misconduct.

For Nigeria, the practical significance lies in extending the use of an existing transparency system without converting a statistical score into an allegation. A workable governance process could operate in five stages. First, BPP or an internal-audit analytics unit would generate periodic risk rankings from the latest sufficiently complete NOCOPO records. Second, auditors would use those rankings to prioritise files for documentary review while also seeing the main factors contributing to the score. Third, the responsible audit or procurement officer would examine the tender file, market conditions, applicable procurement method and any lawful explanation for a single bid before deciding whether further inquiry is warranted. Fourth, escalation to management, the Auditor-General or another competent oversight body would occur only after professional review and documented reasons, with ordinary procedural safeguards preserved for the procuring entity. Fifth, review outcomes would be fed back into model monitoring so that calibration, drift, false positives and override reasons can be assessed over time. In this arrangement, AI supports triage and professional judgment; it does not replace due process, legal interpretation or human responsibility.

The study also has clear scope limitations. The analytical cohort is selected by outcome and date-field availability, so its 9,124 processes should not be treated as representative of all records in the 98,866-process snapshot. The target is single-bid competition, not corruption or legal non-compliance, and it cannot capture legitimate market explanations that are absent from OCDS fields. Finally, the reported model performance is specific to this information set and time window; the rolling-origin results show that discrimination changes across years. These limitations strengthen the case for human review, periodic recalibration and cautious deployment rather than automated enforcement.

V. CONCLUSION AND RECOMMENDATIONS

This study evaluates AI-assisted screening of single-bid competition risk in 9,124 Nigerian federal contracting processes with observable tender-closing dates from 2017 to 2024. Buyer-history and procurement-context variables improve discrimination when compared with simple benchmarks, but the resulting AUC of 0.761 represents useful, moderate screening performance rather than a highly accurate detector. Random forest does not achieve a statistically superior result to logistic regression on the 2023-2024 holdout, while logistic regression is slightly better calibrated. The central lesson is therefore one of disciplined model use rather than technological escalation. Nigerian public-sector accounting and audit analytics

are likely to gain more from reliable records, explicit information timing, transparent benchmarks, calibrated probabilities and continued professional oversight than from selecting the most complex available algorithm.

Five recommendations follow. First, BPP and procuring entities should strengthen data quality, publication lineage and field completeness in NOCOPO because predictive assurance is only as reliable as the underlying information system. Second, public-sector auditors should use risk rankings as triage tools and require documentary review before any adverse inference is drawn. Third, every complex model should be benchmarked against transparent alternatives and retained only when it demonstrates material incremental value in discrimination, calibration or operational usefulness. Fourth, implementation protocols should preserve due process by documenting the reason for escalation, permitting contextual explanations and maintaining a clear human decision maker. Fifth, deployed systems should monitor calibration, false-positive burden and performance drift, especially when procurement conditions or reporting practices change.

Declarations

Ethical Consideration: The study uses public administrative procurement records and does not involve human participants.

Data Availability: The complete data and reproducibility package accompanying this manuscript, including the exact raw BPP/NOCOPO OCDS gzip snapshot, final analytical cohort, model outputs, bootstrap results, holdout predictions, figures, code, codebook and file checksums, is available from Zenodo at <https://doi.org/10.5281/zenodo.22258214>. The original Nigeria BPP/NOCOPO data are publicly distributed through the Open Contracting Partnership Data Registry.

REFERENCES

- Adedire-Ampitan, A. A., Logunleko, S. D., Amusa, B. O., & Onifade, H. O. (2024). Artificial intelligence and risk management of deposit money banks in Nigeria: Empirical evidence from Guaranty Trust Bank. *UMYU Journal of Accounting and Finance Research*, 7(1), 178-193. [https://doi.org/10.61143/umyu-jafr.7\(1\)2024.012](https://doi.org/10.61143/umyu-jafr.7(1)2024.012)
- Alles, M. G., Kogan, A., & Vasarhelyi, M. A. (2008). Putting continuous auditing theory into practice: Lessons from two pilot implementations. *Journal of Information Systems*, 22(2), 195-214.
- Breiman, L. (2001). Random forests. *Machine Learning*, 45, 5-32. <https://doi.org/10.1023/A:1010933404324>
- Bureau of Public Procurement. (2007). Public Procurement Act 2007. Federal Republic of Nigeria.
- Bureau of Public Procurement. (n.d.). Nigeria Open Contracting Portal (NOCOPO). <https://nocopo.bpp.gov.ng/>
- Fazekas, M., & Kocsis, G. (2020). Uncovering high-level corruption: Cross-national objective corruption risk indicators using public procurement data. *British Journal of Political Science*, 50(1), 155-164. <https://doi.org/10.1017/S0007123417000461>
- Fazekas, M., Tóth, I. J., & King, L. P. (2016). An objective corruption risk index using public procurement data. *European Journal on Criminal Policy and Research*, 22(3), 369-397. <https://doi.org/10.1007/s10610-016-9308-z>
- Fedyk, A., Hodson, J., Khimich, N., & Fedyk, T. (2022). Is artificial intelligence improving the audit process? *Review of Accounting Studies*, 27, 938-985. <https://doi.org/10.1007/s11142-022-09697-x>

- Föhr, T. L., Reichelt, V., Marten, K.-U., & Eulerich, M. (2025). A framework for the structured implementation of process mining for audit tasks. *International Journal of Accounting Information Systems*, 56, 100727. <https://doi.org/10.1016/j.accinf.2025.100727>
- Gallego, J., Rivero, G., & Martínez, J. (2021). Preventing rather than punishing: An early warning model of malfeasance in public procurement. *International Journal of Forecasting*, 37(1), 360-377. <https://doi.org/10.1016/j.ijforecast.2020.06.006>
- 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)
- Kokina, J., Blanchette, S., Davenport, T. H., & Pachamanova, D. (2025). Challenges and opportunities for artificial intelligence in auditing: Evidence from the field. *International Journal of Accounting Information Systems*, 56, 100734. <https://doi.org/10.1016/j.accinf.2025.100734>
- Li, Y., & Goel, S. (2025). Artificial intelligence auditability and auditor readiness for auditing artificial intelligence systems. *International Journal of Accounting Information Systems*, 56, 100739. <https://doi.org/10.1016/j.accinf.2025.100739>
- Odiase, O. V., & Uagbale-Ekatah, R. E. (2025). Artificial intelligence and human resources accounting of manufacturing companies in Nigeria. *UMYU Journal of Accounting and Finance Research*, 8(1), 273-283. [https://doi.org/10.61143/umyu-jaf.8\(1\)2025.018](https://doi.org/10.61143/umyu-jaf.8(1)2025.018)
- Open Contracting Partnership. (2026). Nigeria: Bureau of Public Procurement (BPP), Open Contracting Data Registry. <https://data.open-contracting.org/en/publication/64>
- Saito, T., & Rehmsmeier, M. (2015). The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. *PLOS ONE*, 10(3), e0118432. <https://doi.org/10.1371/journal.pone.0118432>
- Stratopoulos, T. C., & Wang, V. X. (2025). Artificial intelligence and accounting research: A framework and agenda. *International Journal of Accounting Information Systems*, 56, 100760. <https://doi.org/10.1016/j.accinf.2025.100760>
- Vasarhelyi, M. A., Kogan, A., & Tuttle, B. M. (2015). Big data in accounting: An overview. *Accounting Horizons*, 29(2), 381-396. <https://doi.org/10.2308/acch-51071>
- Wang, R. Y., & Strong, D. M. (1996). Beyond accuracy: What data quality means to data consumers. *Journal of Management Information Systems*, 12(4), 5-33.