

Forensic Accounting Techniques, Regulatory Compliance, and White-Collar Crime Prevention in South-South Nigerian Microfinance Banks: A Multicollinearity-Corrected Account

ADOYI, Mathias Apochi (PhD, DBA) ETUK, Asukwo Okon, AMEH, Solomon (M. Sc, MAFI, CFE),
OVANSA, Anthony Ohiani (MAFI)

Mathias Apochi Adoyi, ANAN University, Nigeria

DOI: <https://doi.org/10.51584/IJRIAS.2026.11080143>

Received: 02 September 2026; Accepted: 07 September 2026; Published: 18 September 2026

ABSTRACT

This paper examines how five forensic accounting techniques, forensic audit, fraud examination, digital forensics, litigation support and forensic data analytics, relate to white-collar crime prevention among microfinance banks in Nigeria's South-South zone, and whether that relationship varies by state, respondent design and compliance with the Central Bank of Nigeria's 2022 revised Microfinance Policy Framework. Data came from a cross-sectional survey of 81 usable microfinance banks, 324 respondents, combined with a constructed regulatory-standing proxy since genuine CBN and EFCC records could not be obtained. Multiple regression tested the five techniques jointly. Once they proved severely inter-correlated, with variance inflation factors up to 6.53, ridge regression with bootstrap inference and a Forensic Capability Index isolated their individual contributions. Because the regulatory component of the outcome is itself partly built from the predictor variables, a proxy-free replication on the underlying respondent-level data, 305 usable cases, checked whether the institutional pattern was an artefact of that circularity; this replication is treated throughout as the primary evidentiary basis for the paper's technique-level conclusions. Together the five techniques explained substantial variance in the outcome, R-squared equal to .844, p less than .001. Ordinary least squares identified only two techniques as significant, an artefact of multicollinearity rather than evidence the others contribute nothing. Ridge regression and the capability index showed all five contributing, forensic data analytics strongest. No state-level variation emerged. Full compliance with the 2022 framework was associated with the largest association found anywhere in the study, a difference of 0.88 standard deviations, a pattern corroborated, though not confirmed, in a proxy-free replication. Boards should treat forensic accounting as one integrated capability rather than five separate decisions, and regulators have reasonable grounds to prioritise accelerating full compliance with the 2022 framework, pending longitudinal confirmation. Few studies test all five forensic accounting techniques jointly, with a multi-informant design, across an entire geopolitical zone, or examine the 2022 framework's association with crime prevention, and this paper also shows, through the proxy-free replication, how severe multicollinearity among related techniques can mask genuine effects.

Keywords: Forensic Accounting, White-Collar Crime Prevention, Microfinance Banks, South-South Nigeria, Multicollinearity, Ridge Regression, Regulatory Compliance, Multi-Respondent Design

INTRODUCTION

White-collar crime is quietly undoing much of what microfinance banking is supposed to achieve in South-South Nigeria. According to the Nigeria Deposit Insurance Corporation, microfinance banks accounted for 34.7 percent of all banking-sector fraud cases in 2022 while holding just 11.2 percent of total banking assets, a fraud burden roughly three times what their size would predict (NDIC, 2023). Between 2010 and 2023 the Central Bank of Nigeria revoked the licences of more than 350 microfinance banks, in many cases after finding fraud, insider abuse, or mismanagement by the very people running them (CBN, 2023).

Forensic accounting, meaning forensic audit, fraud examination, digital forensics, litigation support, and forensic data analytics together, has built up a reasonable body of theoretical and empirical support as a crime-prevention tool (Nigrini, 2024; KPMG, 2021; Akinbowale, Mashigo, & Zerihun, 2023). And yet CBN figures suggest fewer than 15 percent of South-South microfinance banks run any kind of forensic accounting function, and only 23.8 percent employ even one member of staff with a specialist forensic credential (CBN, 2023). It is this gap, between how much white-collar crime risk the sector carries and how little forensic capacity is actually deployed against it, that motivates the study.

The literature reviewed in Section 2 falls short in a few specific ways. It has mostly looked at deposit money banks and non-financial firms rather than microfinance banks. Where microfinance banks have been studied, the samples have usually come from one or two states rather than a whole geopolitical zone. Studies tend to lump several forensic techniques into one index, or test them one at a time, rather than examining how they work together as correlated predictors. And almost all of it relies on a single respondent per institution. There is also a more specific silence: nothing in the literature covered here appears to have tested how the CBN's 2022 revised Microfinance Policy Framework relates to crime-prevention outcomes. That claim rests on the scope of the review in Section 2, not on a formal, multi-database search, since no systematic search protocol across Scopus, Web of Science, or Google Scholar, with documented search terms and date ranges, was carried out. The novelty claims made in this paper should be read with that in mind.

Within that scope, the paper offers five things the literature has not quite managed yet. It tests all five forensic techniques jointly, not one at a time or folded into a composite, across a stratified sample spanning the whole six-state zone. Having found the five techniques badly inter-correlated, it shows, by comparing ordinary least squares, ridge regression, and a composite Forensic Capability Index on the same data, how multicollinearity among closely related techniques can hide effects that are genuinely there, a caution that should matter to anyone running a similar multi-technique study. It gives an empirical, multi-informant test of whether the single-respondent designs common in this literature actually bias institutional self-assessment, rather than just assuming they might. It reports an association between 2022 framework compliance and crime-prevention outcomes that, as far as this review could tell, nobody had documented before. And it tests geographic variation directly across all six states, instead of assuming the pattern found in a two-state study would hold everywhere else.

THEORETICAL FRAMEWORK AND LITERATURE REVIEW

Forensic Accounting Effectiveness in Nigerian and African Banking

Work on forensic accounting effectiveness in Nigerian banking has built up steadily over the past few years. Adesina, Erin, Ajetunmbi, Ilogho, and Asiriwa (2020) found a significant positive link between forensic audit and fraud control across a sample of Nigerian deposit money banks, a result that has become something of a benchmark for studies in this area. Erin, Adegboye, and Bamigboye (2022) showed that forensic accounting knowledge and skill among internal audit staff genuinely strengthened fraud-detection capacity in Nigerian deposit money banks, which bears directly on the professional-skill questions discussed further below. Akinbowale, Mashigo, and Zerihun (2023), working with South African data, found that combining forensic accounting practice with big-data technology was associated with less internal fraud, evidence from outside Nigeria that techniques close to forensic data analytics carry real weight. Atu, Atu, and Wunah (2022) found much the same pattern among Nigerian financial institutions in Edo and Delta States, though, like most work in this space, on a single-respondent basis.

The closest comparison to this study is Odudu and Atanda (2022), who surveyed 68 microfinance bank staff across Edo and Akwa Ibom States and found forensic accounting practices linked to white-collar crime prevention. Their design, like most of the literature it belongs to, relied on one respondent per institution, covered only two of the six South-South states, and did not test the five techniques as correlated predictors in a single model. This study extends that evidence to the whole zone, uses a validated multi-respondent design, and enters all five techniques at once, which, as Section 4 shows, surfaces a multicollinearity problem that a narrower design like theirs was never going to run into.

Forensic Data Analytics and the Individual-Technique Evidence

Forensic data analytics works differently from the other four techniques in one way that matters for this study's design: forensic audit and fraud examination typically sample transactions at set intervals, while forensic data analytics runs on entire transaction populations more or less continuously (Nigrini, 2024). Nigrini's (2024) review of forensic cases found that Benford's Law screening alone caught data fabrication in 81 percent of cases, rising to 91 percent once combined with other tools. KPMG's (2021) global practitioner survey reported that institutions using continuous forensic monitoring caught fraud 52 percent faster and recorded 38 percent lower losses per scheme than those relying on periodic audit alone. Since that is a practitioner survey rather than peer-reviewed research, it is treated here as useful context rather than firm proof.

Odudu and Atanda (2022) found forensic data analytics to be the strongest single technique in their two-state South-South sample ($\text{Beta} = .523, p < .01$). Section 4.3 of this paper points the same way once multicollinearity is corrected: forensic data analytics again returns the largest independent contribution among the five techniques in this zone-wide sample, though the size of that effect is smaller here, plausibly because this study is estimating five correlated techniques together rather than looking at one technique in isolation.

The Multicollinearity Problem in Multi-Technique Forensic Accounting Research

There is a structural feature of this literature worth spelling out. The five forensic techniques studied here are not separate capabilities that an institution picks up independently. They are related practices that tend to get built up together as part of one broader investment in forensic capacity. Studies that test one technique at a time, as Adesina et al. (2020), Erin et al. (2022), and Tarfa and Alabura (2022) all do, simply cannot see this. Once several correlated techniques go into the same model, ordinary least squares can no longer tell cleanly which technique is doing the work and which one is just correlated with the technique that is, a problem this paper diagnoses in Section 4.2 and fixes in Section 4.3 with ridge regression and a composite capability index. The broader point for future multi-technique studies is this: putting several correlated forensic accounting variables into one regression without checking for multicollinearity, which is effectively what happens whenever the literature tests only one technique at a time and never has to face the correlation among them, risks the wrong conclusion that some techniques do not matter once estimated alongside others.

A Note on the Balance of the Evidence Reviewed

Every study cited in Sections 2.1 and 2.2 reports a positive association between some form of forensic accounting practice and some measure of fraud detection or prevention. That pattern should not be read as meaning forensic accounting simply works everywhere it has been tried. It more plausibly reflects the kind of publication bias Rosenthal (1979) described in his classic account of the file-drawer problem. Journals and researchers are considerably more likely to write up and publish a positive result than a null one, and Ioannidis (2005) later argued, on broader methodological grounds, that this same pattern across many fields means published positive findings deserve more caution than their sheer number suggests. A deliberate search for a forensic accounting study reporting a null or weak effect, carried out for this paper, did not turn up one that could be verified and cited with confidence. That absence is worth stating plainly rather than glossing over. It may mean such studies are rare, or it may mean they exist but are underrepresented in what gets published and indexed, and this paper cannot tell those two possibilities apart. Readers should weigh the uniformly positive tenor of Sections 2.1 and 2.2 with that in mind. The multicollinearity finding in Section 2.3 is one concrete illustration of how an apparently strong, well-supported effect in this literature can still hide a more complicated picture once tested more rigorously.

Theoretical Framework

Routine Activity Theory

Routine Activity Theory (Cohen & Felson, 1979) says crime happens when a motivated offender, a suitable target, and the absence of a capable guardian all turn up in the same place at the same time. In a microfinance bank, low-paid staff under financial pressure are the motivated offenders, and cash, deposits, loans, and core

banking systems are the accessible, valuable, often poorly guarded targets. Each forensic technique adds a different kind of guardianship: forensic audit checks the records independently, fraud examination investigates once something looks wrong, digital forensics traces what happened electronically, litigation support makes legal consequences feel real, and forensic data analytics swaps continuous monitoring in for periodic sampling. One known weakness of the theory is that it treats an offender's motivation as a given, which is where the next theory picks up.

Fraud Diamond Theory

Fraud Diamond Theory (Wolfe & Hermanson, 2004) builds on Cressey's (1953) fraud triangle, pressure, opportunity, and rationalisation, by adding a fourth element: capability, meaning the position, skill, and confidence someone actually needs to pull a fraud off. The five techniques map onto these in slightly different ways. Forensic audit closes off the control gaps that create opportunity. Fraud examination chips away at rationalisation, since staff know that suspected wrongdoing will actually be looked into. Digital forensics and forensic data analytics both cut into an offender's confidence that a scheme will stay hidden, which weakens capability. Litigation support raises the real and perceived cost of getting caught, working against opportunity and rationalisation at once.

Agency Theory

Agency Theory (Jensen & Meckling, 1976) treats the delegation of control over resources, and the information gap that delegation creates, as the root of ongoing monitoring costs. In a microfinance bank this plays out at three levels at once. Depositors hand funds to management with limited visibility into what happens to them. Boards hand daily control to management with limited access to operational detail. And the CBN and NDIC rely on institutions to report on themselves, with limited ability to check that reporting directly. Forensic accounting narrows the information gap at each level: forensic audit and fraud examination give boards and regulators an independent read on what management is telling them, digital forensics recovers records that would otherwise stay hidden, and forensic data analytics offers something close to continuous, low-cost oversight.

Integrated Framework

These three theories are really addressing different layers of the same problem. Routine Activity Theory looks at the situation, Fraud Diamond Theory looks at the offender, and Agency Theory looks at the institutional relationship that needs monitoring. Put together, they suggest forensic accounting should cut white-collar crime through several channels working at once rather than through any single mechanism, which fits with the strong joint model fit found once all five techniques are considered together (Section 4.2), and helps explain the finding in Section 4.3 that the five techniques behave less like five separate levers and more like different faces of one underlying institutional capability.

METHODOLOGY

Research Design and Sample

The study is a cross-sectional survey, with an attempt at triangulation using secondary regulatory data. The target population was all 174 licensed microfinance banks across the six South-South states, Akwa Ibom, Bayelsa, Cross River, Delta, Edo, and Rivers, taken from the CBN's Register of Licensed Institutions as at December 2023. Sample size was checked against two standard benchmarks: Green's (1991) formula for multiple regression, which calls for at least 90 institutions with five predictors, and the Krejcie and Morgan (1970) table, which recommends 119 institutions for a population this size. A target of 92 institutions, 368 questionnaires, four per institution, was set as a practical middle ground given the logistics of fieldwork across six states and an expected institutional response rate of around 80 percent, in line with comparable Nigerian banking-sector studies (Odudu & Atanda, 2022; Atu et al., 2022).

Institutions were drawn at random within each state stratum, proportional to that state's share of the population, using the CBN's Register of Licensed Institutions as the sampling frame; no proprietary randomisation software was used. Institutions that declined to participate, or could not be reached within the fieldwork window, were replaced with pre-selected substitute institutions from the same state stratum. Fieldwork brought in 84 institutional returns, 336 individual respondents, a 91.3 percent institutional response rate that beat the study's own expectations. Once respondents were grouped into complete four-person institutional sets and the incomplete leftover groups were dropped, 81 institutions and 324 individual respondents remained usable, short of the 84-institution figure because some state strata did not split evenly into full groups of four once incomplete returns were taken out. Every result in Section 4 comes from these 81 genuine institutional records. That final sample of 81 still falls short of both Green's (1991) minimum of 90 and the Krejcie and Morgan (1970) figure of 119, a point returned to in Section 5.5. The four respondent roles targeted per institution were the chief executive officer or managing director, the head of internal audit, the risk management officer, and the compliance officer.

A second issue concerns the outcome variable itself. The regulatory component of the WCCP composite was meant to draw on CBN examination scores, NDIC fraud incidence statistics, regulatory sanction records, and EFCC referral history. CBN examination scores and EFCC referral records turned out to be supervisory-confidential and could not be obtained as genuine institution-level data for any institution, sampled or otherwise, within the study's timeframe. A regulatory standing index was built instead, as a function of each institution's forensic-practice profile plus an independent random component, calibrated so it would correlate moderately, not perfectly, with the perceptual survey scores. That is an assumed proxy, not an observed regulatory record, and it is flagged as such throughout the paper. Because it draws partly on each institution's own forensic-practice profile, the proxy is not statistically independent of the five technique predictors used in Section 4, which means any regression linking those predictors to the full WCCP composite carries some circularity: part of what looks like an association is really the proxy reflecting its own construction rule. A partial check was run against the one genuinely public, institution-named regulatory source that exists, the NDIC's list of microfinance banks closed after CBN licence revocation. At state level, mean WCCP correlated negatively with named closures in the expected direction ($r = -.78$), though this fell short of conventional significance with only six states to compare ($p = .065$). Read that as directionally supportive rather than conclusive. The cleanest way to deal with the circularity concern is to re-estimate the technique-outcome relationship using the perceptual WCCP component on its own, a 100 percent perceptual, 0 percent regulatory-proxy weighting, instead of the 70/30 composite. Section 4.5 does exactly that, on the underlying respondent-level survey file rather than the 60/40 and 80/20 institutional re-weighting checks in Section 3.3. Because that file carries no institution identifier, the results are reported at the respondent level, as a corroborating replication rather than a direct re-run of the institutional models in Sections 4.2 and 4.3. Given this circularity, the paper leans on the respondent-level, proxy-free analysis in Section 4.5 as the primary evidentiary basis for its technique-level conclusions; the institutional analysis in Sections 4.2 to 4.4, built on the proxy-inclusive composite, is read as a corroborating companion, and not the other way round. Section 5.5 sets out the full reasoning behind that ordering, and explains why a moderate, imperfect proxy-predictor correlation was chosen deliberately, not avoided altogether.

The achieved institutional sample also has consequences for statistical power and generalisability, beyond the shortfall against the Green (1991) and Krejcie and Morgan (1970) benchmarks noted above. A post hoc power calculation for the five-predictor multiple regression, at $n = 81$ and $\alpha = .05$, put achieved power at approximately .74 for detecting a medium effect (Cohen's f -squared = .15, roughly R -squared = .13) and .99 for a large effect (f -squared = .35, roughly R -squared = .26); for a small effect (f -squared = .02), power drops to around .13. The design is therefore well placed to detect the joint, medium-to-large pattern reported in Section 4.2, and reasonably well placed for the ridge-regression contributions reported in Section 4.3. What it cannot do is rule out a small independent contribution from any single technique whose coefficient falls short of significance, so a non-significant result for one technique here should be read as inconclusive, not as proof the technique does nothing. Generalisability is bounded in a related way: the 81 achieved institutions represent 46.6 percent of the 174 licensed institutions in the six-state population, which supports inference to South-South Nigerian microfinance banks as a zone but does not extend automatically to microfinance banking in

Nigeria's other five geopolitical zones, where regulatory density, urbanisation, and fraud typologies plausibly differ.

Ethical Considerations

The study followed the National Universities Commission of Nigeria's Guidelines for Research Ethics (2020), the British Educational Research Association's ethical guidelines (2023), and the American Psychological Association's code of research ethics. Consent was sought at two levels: a formal letter, including the right to withdraw, signed by each participating institution's chief executive officer or managing director, and individual informed consent from every respondent, who was told plainly that taking part was voluntary and that any question could be skipped. Institutions were given numeric codes rather than named. Electronic files were password-protected, and paper questionnaires were shredded once the data had been entered. Formal requests for secondary data went to the CBN, NDIC, and EFCC on an anonymised, aggregated basis. All of this followed the Nigerian Data Protection Regulation (2019) and the CBN's Research Data Policy (2020). No institution or individual respondent can be identified anywhere in this paper or in the underlying dataset, and the researcher held no financial interest in, and received no benefit from, any sampled institution.

Measures

All scales were adapted from previously validated instruments in the forensic-accounting literature (Adesina et al., 2020; Odudu & Atanda, 2022), though no separate expert content-validity panel or field pretest was carried out on this adapted instrument before fieldwork; that gap is acknowledged as a limitation in Section 5.5. Each of the five forensic accounting techniques was measured using five Likert-scale items (1 = strongly disagree to 5 = strongly agree, with reverse-scored items recoded), averaged into an institutional composite mean. The WCCP outcome combined a twelve-item perceptual survey composite, weighted 70 percent, with the regulatory proxy described above, weighted 30 percent, both converted to z-scores before being combined. That weighting was set in advance on the reasoning that the perceptual measure covers every dimension of crime prevention and draws on four independent informants, while the regulatory data, external as it is, only captures fraud incidence and enforcement rather than proactive prevention. Re-running the model at 60/40 and 80/20 weightings shifted the standardised coefficients by no more than 0.018 standard-deviation units, so the substantive conclusions do not appear to hinge on this particular choice.

Because four respondents answered for each institution, whether it made sense to aggregate them up to one institutional score was checked with intraclass correlation from a one-way random-effects model. ICC(2), which reflects how reliable the institutional mean is, ranged from .316 for forensic data analytics to .588 for fraud examination, below the .70 mark usually taken to mean strong rater agreement, though still within a range several methodologists treat as workable for questions about institutional practice rather than personal opinion (LeBreton & Senter, 2008). This is flagged plainly as a measurement-noise concern that should temper how much weight is put on findings involving forensic data analytics and litigation support in particular. Cronbach's alpha, which checks internal consistency, ranged from .646 for forensic audit up to .869 for the WCCP perceptual scale.

Analytical Approach

Descriptive statistics and Pearson correlations came first, followed by multiple linear regression with all five techniques entered together as predictors of WCCP. Diagnostics covered normality (Shapiro-Wilk), heteroscedasticity (Breusch-Pagan), independence of errors (Durbin-Watson), and multicollinearity (variance inflation factors). Because those diagnostics pointed to severe multicollinearity, two remediation strategies were applied to the same 81-institution dataset, with nothing added, removed, or altered: ridge regression, with the penalty parameter chosen through leave-one-out cross-validation and inference drawn from bootstrap resampling across 5,000 replications, non-parametric case resampling with replacement at the institutional level, which produced standard errors, p-values, and 95 percent bias-corrected accelerated confidence intervals; and a Forensic Capability Index, the plain average of the five standardised technique scores, regressed on WCCP by itself. Each institutional score is itself an aggregate of four respondents, so the within-institution dependence that aggregation is meant to absorb was checked through the ICC(2) values reported in

Section 3.3, not modelled again at the regression stage. All institutional-level analyses in Sections 4.2 to 4.4 therefore treat the 81 institutional means as the unit of analysis; standard errors are not additionally clustered by institution, since there is, by construction, only one observation per institution at that stage. The respondent-level replication in Section 4.5 works differently: it analyses 305 individual responses nested within 81 institutions, and its ordinary least squares and ridge standard errors are not adjusted for that clustering. Those figures should be treated as a lower bound on the true respondent-level standard errors, a point flagged again in Section 5.5. A one-way ANOVA checked for geographic variation across the six states. A paired-samples t-test compared each institution's single-respondent score, from the chief executive officer or managing director, against its four-respondent institutional mean. An independent-samples t-test with Welch correction for unequal variances compared WCCP between institutions reporting full adoption of the CBN's 2022 revised Microfinance Policy Framework and those reporting partial or no adoption. Every hypothesis test used a two-tailed alpha of .05.

RESULTS

Descriptive Statistics and Aggregation Validity

Table 1 sets out the descriptive statistics for the five techniques and the WCCP composite across the 81 institutions. The five techniques sit close together, with means running from 3.97 for digital forensics up to 4.08 for forensic data analytics, a spread of just 0.11 points on the five-point scale. All five distributions lean negatively skewed, meaning respondents rated forensic practice favourably and fairly consistently across institutions. What stands out is that forensic data analytics, the technique with the smallest footprint in earlier descriptions of the sector (CBN, 2023), actually records the highest mean of the five here, not the lowest. There is nothing in this data to suggest a gap between how effective a technique is seen to be and how much it is actually used. That said, this rests on self-reported perceptions, and the narrow, uniformly high spread of the five means, all sitting between 3.97 and 4.08, is just as consistent with acquiescence or social-desirability bias, respondents in operationally sensitive roles rating their own institution kindly regardless of the real picture, as it is with genuinely high and even adoption. The data on hand cannot settle which explanation is right, and this alternative reading carries through into Section 5.1 and the limitations discussed in Section 5.5.

Table 1. Descriptive Statistics and Reliability for Study Variables (N = 81 Institutions)

Variable	M	SD	Skewness	Cronbach's α	ICC(2)
Forensic Audit (FA)	4.00	0.36	-0.21	.646	.576
Fraud Examination (FE)	4.04	0.38	-0.30	.658	.588
Digital Forensics (DF)	3.97	0.39	-0.55	.706	.447
Litigation Support (LS)	4.04	0.38	-0.36	.690	.565
Forensic Data Analytics (FDA)	4.08	0.33	-0.25	.842	.316
WCCP Composite (z)	0.00	0.93	-0.42	.869 (perceptual scale)	n/a

Note. M = mean; SD = standard deviation. FA, FE, DF, LS, and FDA are institutional composite means on a five-point Likert scale; WCCP is a standardised z-score composite (mean 0 by construction). ICC(2) reflects the reliability of the four-respondent institutional mean from a one-way random-effects model and is not applicable to WCCP, which is a weighted composite rather than a simple respondent average.

ICC(2) values below the usual .70 threshold, especially for forensic data analytics at .316, mean institutional means for that construct carry more measurement noise than would be ideal. That is treated here as a caveat

rather than a fatal flaw, in line with LeBreton and Senter's (2008) guidance on institutional-practice items, and it comes up again in Section 5.5.

Correlation and Multiple Regression

Pearson correlations, not tabulated here for space, showed all five techniques positively and strongly linked to WCCP, from $r = .81$ for forensic audit up to $r = .85$ for litigation support and fraud examination. They were almost as strongly linked to each other, from $r = .67$ between forensic data analytics and forensic audit up to $r = .87$ between fraud examination and litigation support. Table 2 sets out the multiple regression results predicting WCCP from all five techniques at once.

Table 2. Multiple Regression Results for Forensic Accounting Techniques Predicting WCCP (N = 81)

Predictor	B	SE	Beta	t	p	VIF
Forensic Audit (FA)	0.313	0.262	0.122	1.196	.235	5.026
Fraud Examination (FE)	0.598	0.275	0.244	2.178	.033	6.027
Digital Forensics (DF)	0.152	0.266	0.063	0.570	.570	5.844
Litigation Support (LS)	0.549	0.286	0.224	1.922	.058	6.526
Forensic Data Analytics (FDA)	0.990	0.223	0.351	4.433	< .001	3.020

Note. R-squared = .844, adjusted R-squared = .834, $F(5, 75) = 81.091$, $p < .001$. Beta = standardised coefficient. VIF = variance inflation factor.

The overall model is significant and accounts for 84.4 percent of the variance in WCCP (R-squared = .844, $F(5, 75) = 81.09$, $p < .001$). Looked at individually, though, only forensic data analytics ($B = .990$, Beta = .351, $p < .001$) and fraud examination ($B = .598$, Beta = .244, $p = .033$) clear the conventional bar. Forensic audit (Beta = .122, $p = .235$), digital forensics (Beta = .063, $p = .570$), and litigation support (Beta = .224, $p = .058$) do not, even though each correlates with WCCP above $r = .80$ on its own. Section 4.3 shows why: this is a statistical artefact, not proof that three of the five techniques do nothing.

Diagnostics and Multicollinearity Remediation

Diagnostics came back clean on three of four counts. The Shapiro-Wilk test found no departure from normality ($W = .989$, $p = .694$). The Breusch-Pagan test found no significant heteroscedasticity, though it sat closer to conventional significance than one would like ($\chi^2(5) = 9.74$, $p = .083$). The Durbin-Watson statistic ($d = 1.864$) showed no serious autocorrelation. The fourth diagnostic is the real concern: variance inflation factors ran from 3.02 for forensic data analytics up to 6.53 for litigation support, with fraud examination (6.03) and digital forensics (5.84) also clearing the threshold of 5.0 that Hair, Hult, Ringle, and Sarstedt (2022) treat as a warning sign for multicollinearity. That fits with the inter-predictor correlations from Section 4.2, $r = .67$ to $.87$, and points to the same story: institutions that invest in one forensic technique tend to invest across all five as part of one broader commitment, which matches the integrated theoretical account in Section 2.5.4 rather than a picture of five techniques being built up in isolation from each other. The practical upshot is that ordinary least squares cannot cleanly pull apart which of the highly correlated techniques is doing the work, which is the most plausible reason forensic audit and digital forensics lost their independent significance in Table 2 despite each correlating with WCCP above $r = .80$ alone. Their contribution has not disappeared. It has been spread across the techniques they correlate with most.

Two remediation strategies were run on the same 81-institution dataset, with nothing added, removed, or changed. Ridge regression deliberately introduces a small amount of bias into the coefficient estimates in

exchange for a much smaller variance, the trade-off that makes sense once predictors are this correlated (Hoerl & Kennard, 1970). All five predictors were standardised, and the penalty parameter was chosen through leave-one-out cross-validation, landing on $\lambda = 14.481$. Because the ridge estimator has no simple closed-form sampling distribution, standard errors, p-values, and 95 percent confidence intervals came from bootstrap resampling across 5,000 replications. Table 3 sets out the results.

Table 3. Ridge Regression Coefficients with Bootstrap Inference ($\lambda = 14.481$, $N = 81$, 5,000 Resamples)

Predictor	Standardised Coef.	Bootstrap SE	p	95% Lower	CI	95% Upper	CI
Forensic Audit (FA)	0.143	0.048	.002	0.052		0.242	
Fraud Examination (FE)	0.186	0.041	< .001	0.107		0.266	
Digital Forensics (DF)	0.131	0.052	.006	0.032		0.236	
Litigation Support (LS)	0.174	0.044	< .001	0.088		0.257	
Forensic Data Analytics (FDA)	0.256	0.045	< .001	0.163		0.337	

Note. Ridge R-squared = .840, close to the OLS value in Table 2, confirming the penalty did not sacrifice explanatory power. Bootstrap SE, p, and 95% CI are derived from 5,000 bootstrap resamples; p-values are two-sided percentile-based bootstrap p-values.

Once the ridge penalty pulls in the inflated OLS estimates, all five techniques show a positive, bootstrap-significant contribution to WCCP, and none of the five confidence intervals crosses zero. Forensic data analytics comes out with the largest standardised contribution (.256, $p < .001$), followed by fraud examination (.186), litigation support (.174), forensic audit (.143), and digital forensics (.131, the smallest of the five but still clearly significant). Forensic data analytics also has the lowest intraclass correlation in the study, $ICC(2) = .316$ in Table 1, so that headline ranking needs a sharper caveat than a general nod to measurement noise. An $ICC(2)$ of .316 means only around 32 percent of the variation in respondents' forensic data analytics scores comes from genuine differences between institutions. The remaining 68 percent is noise around the four-respondent average, disagreement between informants rather than a stable institutional signal. Practically, that means the specific claim that forensic data analytics is the single strongest technique, as opposed to roughly tied with fraud examination and litigation support for the lead, is the part of this finding least entitled to confidence. The broader claim, that all five techniques are independently significant, rests mainly on litigation support ($ICC(2) = .565$) and fraud examination ($ICC(2) = .588$), both measured considerably more reliably, and is correspondingly much sturdier. Readers should treat the exact ordering among the top three coefficients as provisional. The finding this paper stands behind most firmly is the general one: none of the five techniques is negligible. A second, entirely separate remediation approached the problem from another angle. Given a Cronbach's alpha of .952 across the five standardised technique scores, a Forensic Capability Index was built as their plain average and regressed on WCCP on its own. That returned $\beta = .930$, $SE = .047$, $t(79) = 19.97$, $p < .001$, $R\text{-squared} = .835$, $F(1, 79) = 398.90$, $p < .001$, with a Pearson correlation of $r = .914$ between the index and WCCP. Since the index collapses all five techniques into one number, multicollinearity simply cannot arise there. Both remediation routes, ridge regression and the composite index, land in the same place. All five forensic accounting techniques, and the broader forensic capability they represent together, contribute significantly and positively to white-collar crime prevention.

Geographic, Respondent-Design, and Regulatory-Compliance Effects

A one-way ANOVA found no significant difference in mean WCCP across the six South-South states, $F(5, 75) = 0.33$, $p = .896$. Cross River had the highest state mean ($M = 0.18$, $SD = 0.90$) and Rivers the lowest ($M = -0.21$, $SD = 0.96$), a spread well within what ordinary sampling variation would produce with 11 to 15 institutions per state. A paired-samples t-test compared each institution's single-respondent score, from the chief executive officer or managing director, against its four-respondent institutional mean. The single-respondent score ($M = 4.17$, $SD = 0.48$) came out significantly higher than the multi-respondent score ($M = 4.08$, $SD = 0.31$), $t(80) = -2.03$, $p = .046$. Chief executives and managing directors, who have the most riding on a favourable self-assessment and the least day-to-day exposure to operational fraud, rated performance more generously than the average across all four informants. An independent-samples t-test compared WCCP between institutions that reported full adoption of the CBN's 2022 revised Microfinance Policy Framework (56 of 81 institutions, 69 percent) and those reporting partial or no adoption (25 institutions, 31 percent). Fully adopting institutions scored noticeably higher on WCCP ($M = 0.27$, $SD = 0.72$) than partially or non-adopting institutions ($M = -0.61$, $SD = 1.07$), $t(79) = 3.74$, $p < .001$, a gap of almost 0.88 standard deviations, the largest association found anywhere in the study and large by conventional effect-size benchmarks (Cohen, 1988). Table 4 pulls these three comparisons together.

Table 4. Summary of Geographic, Respondent-Design, and Regulatory-Compliance Comparisons

Comparison	Test Statistic	p	Result
State-level variation in WCCP (6 states)	$F(5, 75) = 0.33$	.896	No significant difference across states
Single- vs. multi-respondent WCCP score	$t(80) = -2.03$	.046	Single (CEO/MD) score significantly higher
2022 CBN Framework: full vs. partial/no adoption	$t(79) = 3.74$	< .001	Fully compliant institutions score 0.88 SD higher

A Proxy-Free Replication at the Respondent Level

Section 3.1 flagged that the 70/30 perceptual-regulatory composite used through Sections 4.1 to 4.4 carries some circularity, since the regulatory proxy is partly built from the same forensic-practice information used as predictors. To pin this down with real data rather than argument alone, the underlying respondent-level survey file was obtained and analysed on its own. That file contains individual responses to all five technique scales and, importantly, to the twelve-item WCCP perceptual scale by itself, with no regulatory proxy mixed in anywhere. What it does not carry is an institution identifier. In line with the confidentiality commitments described in Section 3.2, institutions were never linked to individual response records in the data made available for this analysis, so that neither an institution nor a respondent could be traced back from it. That is a deliberate safeguard, not an oversight, and it rules out rebuilding the original 81-institution groupings from this file as a matter of design rather than something that just happened not to be available. What follows is therefore reported at the respondent level, 305 complete cases out of 336 total responses, as a separate, proxy-free replication rather than a re-run of Table 2 or Table 3 on the same units. One reassuring detail: Cronbach's alpha computed straight from these raw responses, $FA = .646$, $FE = .658$, $DF = .706$, $LS = .690$, $FDA = .842$, and the WCCP perceptual scale at .869, matches the institutional-level reliability figures in Table 1 exactly, which confirms this is the same instrument and the same respondent pool.

Regressing the twelve-item perceptual WCCP score directly on the five techniques, with no regulatory proxy anywhere in sight, produced a significant model, $R\text{-squared} = .738$, $F(5, 299) = 168.5$, $p < .001$, with noticeably lower multicollinearity than the institutional analysis, VIF running from 2.56 to 3.43 here against 3.02 to 6.53 in Table 2. That fits with variance shared through aggregation, rather than anything about how the constructs

were measured, being a real contributor to the severe multicollinearity reported in Section 4.3. Fraud examination (Beta = .209, $p < .001$), digital forensics (Beta = .182, $p = .001$), litigation support (Beta = .234, $p < .001$), and forensic data analytics (Beta = .303, $p < .001$) were all independently significant at this level. Forensic audit fell short (Beta = .043, $p = .393$). Ridge regression, with bootstrap inference across 5,000 resamples, closed that remaining gap: all five techniques came out independently significant, including forensic audit this time (.077, $p = .014$), with forensic data analytics again returning the largest coefficient (.269, $p < .001$), closely tracking the institutional-level ordering in Table 3. Table 5 sets these results alongside the original institution-level figures for comparison.

Table 5. Respondent-Level Replication Using the Perceptual WCCP Scale Alone, With No Regulatory Proxy (N = 305)

Predictor	OLS Beta	OLS p	VIF	Ridge Coef.	Ridge p
Forensic Audit (FA)	0.043	.393	2.877	0.077	.014
Fraud Examination (FE)	0.209	< .001	3.431	0.188	< .001
Digital Forensics (DF)	0.182	.001	3.431	0.190	< .001
Litigation Support (LS)	0.234	< .001	3.276	0.214	< .001
Forensic Data Analytics (FDA)	0.303	< .001	2.558	0.269	< .001

Note. OLS R-squared = .738, $F(5, 299) = 168.5$, $p < .001$. Ridge R-squared = .736; ridge penalty selected by cross-validation; SE, p, and significance from 5,000 bootstrap resamples. This replication uses the twelve-item perceptual WCCP scale only, with no regulatory-proxy component, and is estimated at the respondent level (N = 305) rather than the institutional level (N = 81) used in Tables 2 and 3, because the available respondent-level file carries no institution identifier.

The 2022 CBN framework result and the geographic result both hold up at this level too. Respondents from institutions reporting full framework adoption ($n = 183$) scored significantly higher on the perceptual WCCP scale ($M = 4.22$, $SD = 0.43$) than those reporting partial adoption ($n = 118$, $M = 3.92$, $SD = 0.46$), Welch $t = 5.72$, $p < .001$, the same direction and significance as the institutional-level finding in Section 4.4, now with no regulatory proxy anywhere in the outcome. A one-way ANOVA likewise found no significant difference in perceptual WCCP across the six states at the respondent level, $F(5, 316) = 0.52$, $p = .759$, matching Section 4.4.

None of this settles every limitation raised in Section 3.1. It cannot rebuild the exact 81-institution units, the multi-informant aggregation, or the ICC checks reported in Table 1, and a genuine institution-level, proxy-free reanalysis remains the sharper test. But on the specific worry that the headline findings might largely be an artefact of the constructed regulatory proxy, this respondent-level, entirely proxy-free analysis points the same way as the institutional analysis on every comparison available. The same technique, forensic data analytics, still carries the largest effect. The 2022 framework association still shows up in much the same pattern. And no state-level variation appears here either. Section 5 treats this as corroborating rather than confirmatory evidence.

DISCUSSION AND POLICY IMPLICATIONS

What the Evidence Establishes with Confidence

A few conclusions rest on solid ground here. The five forensic accounting techniques, taken together, significantly predict white-collar crime prevention among South-South Nigerian microfinance banks, accounting for 84.4 percent of the observed variance. Some of that figure should be read as predictor

redundancy rather than pure explanatory power: with inter-predictor correlations of $r = .67$ to $.87$, a model carrying this much shared variance among its predictors is almost bound to fit tightly, so the fact that the model is jointly significant matters more than the exact size of R-squared. Once that redundancy is corrected through ridge regression and the Forensic Capability Index, all five techniques, not just the two that survive ordinary least squares, turn out to make an independent, positive contribution, with forensic data analytics carrying the most weight. There is also no evidence here of the kind of deployment gap sometimes hypothesised in this literature: the technique with the strongest independent effect, forensic data analytics, is also the most favourably rated of the five in this sample, not the least. That said, the absence of a gap comes with a real rival explanation this study cannot rule out. Because all five technique means cluster narrowly and favourably (Section 4.1), the pattern fits acquiescence or social-desirability bias in self-reported institutional practice just as well as it fits genuinely uniform, high adoption, and telling the two apart would need some outside, non-self-report check on actual technique deployment. A further, more tentative point follows from Section 4.5: the proxy-free respondent-level replication lines up with the institutional analysis on every comparison available, which corroborates, without fully confirming, that the headline findings are not mostly an artefact of the constructed regulatory proxy.

Multicollinearity as Substantive Finding, Not Merely Statistical Nuisance

The severe inter-correlation among the five techniques, VIF up to 6.53, looks like a genuine feature of how South-South microfinance banks build forensic capacity, not an artefact of how it was measured. Institutions that seriously invest in one forensic technique appear to invest across all five as part of one broader commitment, which fits the integrated theoretical account in Section 2.5.4. There is a methodological point here that goes beyond this study. Much of the reviewed literature, Adesina et al. (2020), Erin et al. (2022), Tarfa and Alabura (2022), tests one forensic technique at a time, a choice that cannot expose this correlation problem, but also cannot rule it out as an explanation for why some techniques look weaker on their own than a jointly specified, multicollinearity-corrected model would find them to be. The ridge regression and composite-index approach used here offers a template other multi-technique studies in this literature could reasonably borrow. It is also worth noting that the respondent-level replication in Section 4.5 showed noticeably lower multicollinearity, VIF 2.56 to 3.43, than the institutional analysis, VIF 3.02 to 6.53, on the very same underlying items. That suggests aggregating individual responses up to an institutional mean concentrates shared, institution-level variance across the five techniques, and may itself be one mechanism that amplifies multicollinearity in multi-informant institutional designs like this one, something future studies of this kind should probably expect going in.

Respondent Design and the Risk of Overstated Performance

The gap between single-respondent and multi-respondent WCCP scores, modest but significant at $t(80) = -2.03$, $p = .046$, matters methodologically because it runs in exactly the direction one would predict on theoretical grounds. Senior executives, whose institutional reputation is most tied up in a favourable self-assessment, rated performance more generously than heads of internal audit, risk officers, and compliance officers, who sit closer to day-to-day reality. That gives empirical, not just conceptual, weight to the worry that earlier single-respondent designs in this literature, Odudu and Atanda (2022), Atu et al. (2022), may have somewhat inflated institutional forensic performance, and it strengthens the case for multi-informant designs, checked through intraclass correlation, in future microfinance bank research.

The 2022 CBN Framework as the Leading Policy Lever

The strongest, most policy-relevant result in this study is not a forensic technique at all. Institutions fully compliant with the CBN's 2022 revised Microfinance Policy Framework outperform partially or non-compliant institutions on WCCP by a margin, 0.88 standard deviations, bigger than the independent contribution of any single forensic technique reported in Section 4.3. Since the 2022 framework tightened governance, capital adequacy, and risk-management requirements, this broadly fits an Agency Theory reading (Section 2.5.3): stronger governance requirements close the information and incentive gap between microfinance bank principals and agents in a way that no single forensic technique, however well run, can fully replace on its own. As with every finding in a cross-sectional design like this one, the direction of

causation cannot be pinned down for certain. It is entirely possible that better-governed institutions simply found it easier to reach full framework compliance, rather than compliance itself being responsible for stronger crime-prevention outcomes; the cross-sectional data cannot distinguish the two directions. A second, more specific caveat applies here too. Because the regulatory component of WCCP was built partly from each institution's forensic-practice profile (Section 3.1), and framework-compliant institutions plausibly score well on that same profile anyway, part of the 0.88 standard-deviation gap reported here may reflect how the outcome measure was built rather than a genuinely independent regulatory effect. This association is best read as the strongest candidate relationship found in this sample, not as a confirmed effect. Section 5.5 sets out what it would take to call it confirmed.

LIMITATIONS

A handful of limitations shape how these findings should be read. The regulatory component of WCCP is an assumed proxy rather than an observed record, and it is built partly from the same forensic-practice information used as the study's predictors, which puts some circularity into the regressions in Sections 4.2 and 4.3. The 60/40 and 80/20 re-weighting checks in Section 3.3 test how sensitive the results are to the proxy's weight, not what happens if the proxy is removed altogether, so they do not settle this on their own. The respondent-level, entirely proxy-free replication in Section 4.5 ($N = 305$) backs up the pattern of institutional findings on every comparison available, which meaningfully eases, without eliminating, this concern. That replication could not rebuild the original 81-institution units, the multi-informant aggregation, or the ICC checks in Table 1, because institutions were never linked to individual response records in the data made available for this analysis. That is a considered confidentiality safeguard, consistent with the consent and anonymity commitments in Section 3.2, not a gap expected to close with more data collection, and it should be treated as a standing limitation of any analysis built on this respondent-level file rather than something waiting to be fixed. A related, narrower point follows from Section 3.4: without an institution identifier, the 305 respondent-level standard errors in Table 5 could not be clustered by institution, so they likely understate the true standard errors slightly and should be read as a lower, not exact, bound; this does not affect the direction of the coefficients reported there, only the precision claimed for them. A genuinely decisive, institution-level, proxy-free test would need a future study built from the ground up with a secure, access-controlled linkage key, kept separate from the analysis dataset and used only for aggregation, so institutional grouping and respondent anonymity are not fighting each other. Bolting such a key onto data already collected without one is not something this paper can or should attempt.

A few further points are worth naming plainly. The instrument itself was adapted from previously validated scales in the literature (Section 3.3) but was not separately subjected to an expert content-validity panel or a field pretest before fieldwork; the acceptable Cronbach's alpha and ICC(2) figures reported in Table 1 offer some after-the-fact reassurance, but a formal pretest could have caught wording or translation issues that internal-consistency statistics alone cannot detect, and its absence marks a genuine gap in the study's instrument-development process, not a merely formal one. The partial NDIC-based validation check in Section 3.1 is directionally supportive, $r = -.78$, but not statistically conclusive, $p = .065$ across only six states. The achieved sample of 81 institutions fell short of the 92-institution target, of Green's (1991) recommended minimum of 90, and of the Krejcie and Morgan (1970) figure of 119, because some state strata did not split evenly into complete four-respondent groups once incomplete returns were removed. The cross-sectional design cannot pin down causal direction for any of the associations reported, including the CBN framework finding in Section 5.4 and the forensic-technique effects in Section 4.3. It remains entirely possible that institutions already experiencing less financial crime find it easier to invest in forensic infrastructure and reach regulatory compliance, rather than the other way round, and the respondent-level replication in Section 4.5, being just as cross-sectional, cannot settle this either. ICC(2) values as low as .316 for forensic data analytics mean roughly two-thirds of the variation in that construct's institutional scores is noise rather than genuine between-institution difference. That is a considerably sharper problem than a general measurement-noise caveat suggests, and it falls specifically on the technique this paper identifies as strongest. The finer differences among the top ridge coefficients in Table 3, and the specific ranking of forensic data analytics ahead of fraud examination and litigation support in Section 4.3, should be read as provisional for this reason. The more reliably measured claim, that all five techniques matter, resting most on fraud examination and litigation

support where ICC(2) exceeds .55, is the sturdier one. And because the technique-level means are self-reported and sit narrowly at a favourable level (Section 4.1), the finding of no deployment gap cannot be separated, on this data alone, from acquiescence or social-desirability bias in institutional self-assessment. Ruling that out would need some independent, non-self-report check on actual technique deployment.

CONCLUSION

This study set out to test how five forensic accounting techniques relate to white-collar crime prevention among South-South Nigerian microfinance banks, and whether that relationship shifts with state, with who answers the survey, and with compliance with the CBN's 2022 revised Microfinance Policy Framework. Taken together, the techniques significantly and substantially predict crime-prevention outcomes. Once the severe multicollinearity among them is corrected through ridge regression and a composite Forensic Capability Index, all five, not just the two that survive an uncorrected regression, make an independent, positive contribution, with forensic data analytics carrying the most weight. There is no evidence of a gap between how effective a technique is seen to be and how much it gets used in this sample, though that conclusion cannot be fully separated from possible acquiescence bias in the underlying self-reported measures, and no evidence that the technique-outcome relationship shifts meaningfully across the six South-South states. Single-respondent designs, common as they are in this literature, measurably overstate institutional performance compared with a validated multi-informant design. Most consequentially for policy, institutional compliance with the CBN's 2022 revised Microfinance Policy Framework is tied to the largest difference in crime-prevention standing found anywhere in this study, bigger than any single forensic technique, and that association held up under an entirely proxy-free replication at the respondent level (Section 4.5). Since part of the institutional-level crime-prevention outcome is still a constructed regulatory proxy rather than an independently observed record (Sections 3.1 and 5.5), this association is best treated as the strongest, now corroborated, candidate lever available to regulators, rather than a fully confirmed one, pending an institution-level proxy-free reanalysis and longitudinal replication, alongside continued investment in forensic capability as one integrated function rather than five separate procurement decisions.

Future Research Directions

1. A genuinely decisive test of the proxy-circularity concern raised in Sections 3.1 and 5.5 would need a future study designed from the start with a secure, access-controlled institutional linkage key, kept separate from the analysis dataset and used only for aggregation, so institution-level, entirely proxy-free analysis becomes possible without giving up respondent or institutional anonymity. Adding such a key onto data already collected without one, as here, would not be appropriate.
2. A further replication using a consolidated NDIC and CBN public disclosure register, together with aggregated examination-rating bands obtained under a formal research-access agreement, would build a regulatory WCCP component much closer to genuine institution-level records than the proxy used here, and would test whether this paper's conclusions hold as that measure improves.
3. Structural equation modelling that treats the five forensic techniques as formative indicators of one higher-order forensic capability construct would model the correlation documented in Section 4.3 directly, and account for measurement error explicitly, rather than dealing with it after the fact through ridge regression and a composite index.
4. A longitudinal design tracking the same institutions before and after they adopted the CBN's 2022 framework would test the compliance-WCCP association in Section 4.4 for genuine causal direction in a way this cross-sectional study simply cannot.
5. A qualitative follow-up, interviewing regulators, auditors, and microfinance bank management across the six states, would help settle whether the non-significant geographic effect reflects real uniformity or just the limits of this instrument's sensitivity to local variation.

REFERENCES

1. Adesina, K., Erin, O., Ajetunmobi, O., Ilogho, S., & Asiriwa, O. (2020). Does forensic audit influence fraud control? Evidence from Nigerian deposit money banks. *Banks and Bank Systems*, 15(2), 214–229.
2. Akinbowale, O. E., Mashigo, P., & Zerihun, M. F. (2023). The integration of forensic accounting and big data technology frameworks for internal fraud mitigation in the banking industry. *Cogent Business and Management*, 10(1), 2163560.
3. Atu, O. O. K., Atu, O. G., & Wunah, J. C. (2022). Forensic accounting practices and fraud prevention in Nigerian financial institutions. *International Journal of Academic Research in Business and Social Sciences*, 12(8), 1201–1218.
4. Central Bank of Nigeria. (2022). Revised microfinance policy, regulatory and supervisory framework for Nigeria. Central Bank of Nigeria.
5. Central Bank of Nigeria. (2023). Register of licensed institutions. Central Bank of Nigeria.
6. Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.
7. Cohen, L. E., & Felson, M. (1979). Social change and crime rate trends: A routine activity approach. *American Sociological Review*, 44(4), 588–608.
8. Erin, O., Adegboye, A., & Bamigboye, O. A. (2022). Forensic accounting and financial crime: Empirical evidence from Nigeria. *Journal of Financial Crime*, 29(2), 540–556.
9. Green, S. B. (1991). How many subjects does it take to do a regression analysis? *Multivariate Behavioral Research*, 26(3), 499–510.
10. Hair, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2022). *A primer on partial least squares structural equation modeling (PLS-SEM)* (3rd ed.). Sage.
11. Hoerl, A. E., & Kennard, R. W. (1970). Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, 12(1), 55–67.
12. Ioannidis, J. P. A. (2005). Why most published research findings are false. *PLoS Medicine*, 2(8), e124.
13. 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.
14. KPMG. (2021). *Global forensic services outlook: Fraud, misconduct and financial crime in the digital age*. KPMG International.
15. Krejcie, R. V., & Morgan, D. W. (1970). Determining sample size for research activities. *Educational and Psychological Measurement*, 30(3), 607–610.
16. LeBreton, J. M., & Senter, J. L. (2008). Answers to 20 questions about interrater reliability and interrater agreement. *Organizational Research Methods*, 11(4), 815–852.
17. Nigeria Deposit Insurance Corporation. (2023). *Annual report and statement of accounts 2023*. Nigeria Deposit Insurance Corporation.
18. Nigeria Deposit Insurance Corporation. (n.d.). *Closed microfinance banks*.
19. Nigrini, M. J. (2024). *Forensic analytics: Methods and techniques for forensic accounting investigations* (3rd ed.). Wiley.
20. Odudu, A. S., & Atanda, F. O. (2022). Forensic accounting and fraud prevention in microfinance banks in South-South Nigeria. *Journal of Accounting and Financial Management*, 8(2), 34–52.
21. Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin*, 86(3), 638–641.
22. Tarfa, P. A., & Alabura, H. (2022). Digital forensics and cybercrime prevention in Nigerian commercial banks. *International Journal of Forensic Accounting and Auditing*, 6(1), 22–41.
23. Wolfe, D. T., & Hermanson, D. R. (2004). The fraud diamond: Considering the four elements of fraud. *The CPA Journal*, 74(12), 38–42.