# The Scale–Technology Paradox: Technological Progress and Productivity Decline in Nigerian Banking*

Abiodun M. Musbaudeen † Subal C. Kumbhakar‡

**Abstract**

Nigeria's banking sector has undergone substantial consolidation and technological modernization, yet bank productivity has continued to decline. This study examines this scale–technology paradox using annual data for 17 commercial banks over 2000–2023. We estimate a multi-output input distance function using a control-function approach to distinguish productivity change from technical change and examine the roles of scale economies and input costs. Bank productivity declined by 3.5% per year on average (3.1% at the median), with larger and internationally licensed banks experiencing steeper declines. Technical change was cost-reducing, while returns to scale averaged 1.46, indicating substantial unexploited scale economies. Lending was the main source of cost pressure. The findings show that technological progress and scale economies did not translate into productivity growth because scale economies remained underexploited and lending costs persisted. Strengthening credit information, collateral enforcement, bad-loan resolution, and lending efficiency is therefore central to translating technological progress and scale economies into sustained productivity growth.

Keywords: Bank productivity; cost efficiency; scale economies; technical change; Nigerian banking sector

JEL Classification No.: C33, D24, G21

*We specially thank the editor and two anonymous reviewers for their valuable comments and suggestions that improve the quality of the paper. We also thank David Slichter, Solomon Polachek, Yingjun Su, Ayodele Shittu, Ivan Korolev, and JoonHwan Cho for their valuable comments on earlier drafts of this paper. Finally, we thank the participants of the 2025 North America Productivity Workshop (NAPW) for their constructive feedback. All remaining errors are our own.

†Department of Economics, State University of New York, Binghamton, NY, USA, amusbau1@binghamton.edu

‡Corresponding author, Department of Economics, State University of New York, Binghamton, NY, USA kkar@binghamton.edu

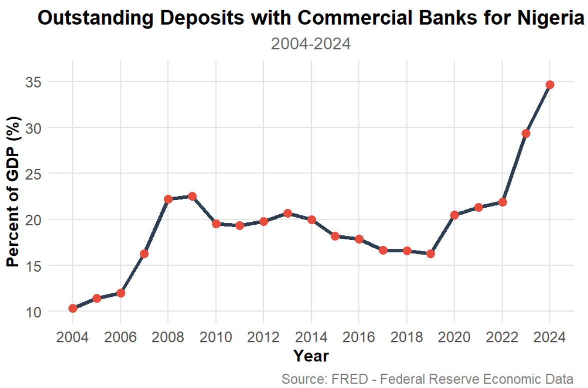
## Highlights

- The study uses annual-report data for 17 Nigerian banks from 2000 to 2023.
- A multi-output distance function separates productivity from technical change.
- Bank productivity fell by 3.5% per year despite cost-reducing technology.
- Loans are the main cost driver, with a mean output elasticity of 0.501.
- Unrealized scale economies help explain weak productivity performance.

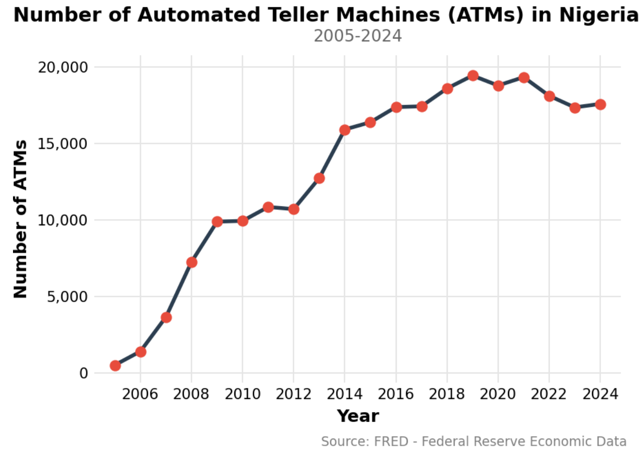
# 1 Introduction

Productivity in the banking sector plays a central role in shaping real economic outcomes because banks allocate financial resources across firms, households, and industries. Through intermediation, banks transform mobilized deposits into credit and other financial assets, thereby influencing investment, consumption, and aggregate output. A more productive banking system improves the efficiency of this allocation process, lowers the cost of financial services, and supports macroeconomic stability and long-run growth. Nigeria's banking sector has undergone extensive structural transformation since the return to democratic governance in 1999. As shown in Figure 1, financial depth expanded markedly: outstanding bank deposits increased from 10.4% of GDP in 2004 to 21.9% in 2022 and reached 34.6% in 2024 (International Monetary Fund, 2026d). Delivery networks also expanded, with ATMs rising from 532 in 2005 to 18,136 in 2022 (International Monetary Fund, 2026a). Branch networks initially grew, peaking at 5,810 in 2011, before consolidating to 4,522 in 2022 (International Monetary Fund, 2026b).

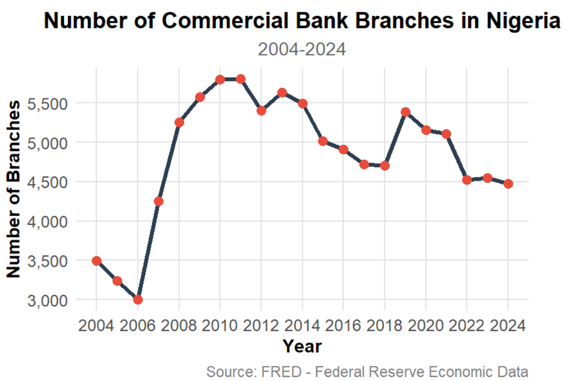
Regulatory reforms, particularly increases in minimum capital requirements, precipitated a major consolidation of the banking industry. The number of commercial banks declined from 89 in 2004 to 25 in 2005 and has remained broadly stable at about 26 in 2024 (International Monetary Fund, 2026c). These simultaneous developments, through balance-sheet expansion, network restructuring, and changes in market structure, reflect the sector's adjustment to sizable technological and regulatory shocks and underscore the need for a systematic assessment of productivity dynamics over 2000–2023. Despite this expansion, substantial inefficiencies persist. Rising non-performing loans, resource misallocation, and governance failures have periodically undermined financial stability, most notably during the 2009 crisis (Sanusi, 2010). Weak credit information systems, limited collateral enforcement, and adverse selection continue to elevate credit risk and operational costs. These structural frictions make productivity measurement particularly important for understanding bank behavior and systemic vulnerability.



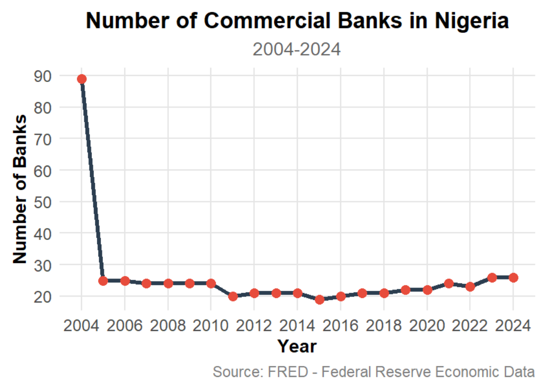

Panel A: Outstanding Deposits (% of GDP)




Panel B: Number of ATMs




Panel C: Number of Bank Branches




Panel D: Number of Commercial Banks

Figure 1: Evolution of Nigeria's Banking Sector Infrastructure and Financial Depth, 2004–2024

In this paper, we address two important questions. First, how has banking productivity evolved since Nigeria's structural reforms began? Second, given increased deposits, greater concentration, and expanded financial infrastructure, are banks realizing productivity gains consistent with economies of scale, or do rising costs signal persistent inefficiency? These questions are especially salient in the Nigerian context, where banks face a dual challenge of relatively low credit supply and a high incidence of delinquent loans. Together, these challenges point to informational frictions and heterogeneity in managerial quality, both of which shape input choices and cost structures. We therefore define productivity, for the purposes of this analysis, as the ability to transform inputs — deposits and borrowings, labor, and capital — into performing loans and investments.

We model banks as multi-output firms and estimate an input distance function (IDF) using a control function approach grounded in cost minimization behavior. The IDF framework is particularly well-suited for this application because it accommodates multiple outputs, does not require input price data (unavailable in our dataset), and facilitates identification of both productivity and technical change.

Although control function approaches in the literature (Olley & Pakes, 1996; Levinsohn & Petrin, 2003; Ackerberg et al., 2015) address endogeneity concerns, they typically assume a single endogenous output and rely on optimal input choices derived under profit maximization to proxy for unobserved productivity. This assumption is unsuitable for banking, where outputs such as loans and investments are substantially influenced by exogenous factors—including monetary policy, regulatory requirements, and macroeconomic credit demand—but are beyond an individual bank's control. In other words, output variables are not endogenous (decision) variables for a bank (Das & Kumbhakar, 2012; Wang et al., 2025). Consequently, the IDF framework is suitable for analyzing multiple outputs in the banking sector.

Using an unbalanced panel of 17 Nigerian commercial banks over 2000–2023, we find an average productivity change of $-3.5\%$ per annum and a median of $-3.1\%$, reflecting a sustained decline in banking productivity over the sample period. Loans are costlier to produce than investments, with a mean output elasticity of 0.501 for loans versus 0.187 for investments. The yearly mean RTS exceeds unity throughout the period, with a sample-wide mean of 1.46 and a median of 1.47, indicating persistent unrealized economies of scale. Technical change indicates cost reductions throughout the period, with a mean of $-0.035$. The coexistence of technical progress and declining productivity indicates that scale sub-optimality and related operating frictions remain important sources of cost inefficiency. Model-based robustness checks preserve the baseline decline.

This paper makes four contributions. First, we model Nigerian banks as multi-product firms, addressing limitations of single-output studies. Second, we provide the first application of a control function–based input distance function estimator to banking productivity in a developing economy. Third, our approach separates productivity from technical change, yielding a clearer understanding of cost dynamics. Fourth, by documenting productivity change, scale economies, and technical change, we provide evidence relevant for policy discussions on consolidation, technological upgrading, and cost-efficiency reforms.

The remainder of the paper proceeds as follows. Section 2 reviews the literature. Section 3 provides institutional background on the Nigerian banking sector. Section 4 outlines the bank production model. Section 5 describes the data. Section 6 presents the results and robustness checks. Section 7 discusses the findings and policy implications, while Section 8 concludes.

## 2 Related Literature

The literature on banking performance has progressed from the use of simple financial ratios to frontier-based methods that characterize how banks transform funding and other inputs into loans, investments, and financial services. These studies show that bank productivity depends on technology, scale, credit risk, and the organization of the intermediation process. Most empirical studies evaluate technical, cost, allocative, or profit efficiency using either Data Envelopment Analysis (DEA) or Stochastic Frontier Analysis (SFA) (Berger et al., 1987; Berger, 2007; Delis & Tsionas, 2009; Barros et al., 2012; Asmild

& Matthews, 2012; Das & Kumbhakar, 2012). Survey evidence confirms the prominence of these approaches. Berger (2007) classifies more than 100 studies into cross-country comparisons, national-frontier analyses, and comparisons of domestic and foreign banks, while Fethi & Pasiouras (2010) review 196 studies and document the dominance of DEA alongside the growing use of artificial-intelligence methods.

This large literature is geographically concentrated in the United States, Europe, and parts of Asia. As Berger (2007) observes, this concentration largely reflects differences in data availability rather than the relative economic importance of banking systems across regions. Evidence from developing economies, particularly sub-Saharan Africa, remains comparatively limited (Berger et al., 1987; Berger, 2007; Das & Kumbhakar, 2012). Notably, the major surveys do not cover Nigerian banking studies, illustrating the limited representation of the country in the broader banking-efficiency literature.

Existing Nigerian studies primarily estimate static efficiency scores using DEA or SFA (Zhao & Murinde, 2011; Osuagwu et al., 2018). Although these studies provide useful evidence on differences in bank performance, they generally offer limited information about the evolution and sources of productivity over time. In particular, static efficiency estimates do not necessarily reveal whether improvements in performance arise from banks moving closer to the production frontier or from the frontier itself shifting because of technological progress. Consequently, relatively little is known about the long-run dynamics of productivity and technical change in Nigerian banking.

Recent frontier studies highlight four issues that are central to our analysis: the measurement and decomposition of productivity, the multi-output nature of banking, input endogeneity, and the role of the operating environment and asset quality.

The first issue concerns the measurement of productivity and technological change. Tsionas & Mallick (2019) develop a Bayesian semiparametric stochastic-frontier model with latent dynamic productivity, illustrating the importance of accounting for the evolution and persistence of productivity over time. In banking, Mallick et al. (2020) use a time-dependent nonparametric approach to distinguish movements in the technological frontier from banks' movements toward it. They find that efficiency convergence occurs in clubs whose composition is partly shaped by bank ownership. These studies demonstrate that observed performance changes may reflect technological change, efficiency catch-up, or persistent productivity differences and therefore motivate our separation of productivity change from shifts in the production frontier.

The second issue concerns the multi-output nature of banking. Banks jointly produce loans, investments, and other financial services using common inputs, so treating them as single-output firms may misrepresent their production technology and distort estimates of scale and productivity. Multi-output approaches are therefore particularly appropriate for analyzing financial institutions. Studies such as Das & Kumbhakar (2012) for India and Wang et al. (2025) for the United States demonstrate the usefulness of multi-output frameworks, but this approach has received limited application in the Nigerian banking literature.

The third issue is input endogeneity. Banks select labor, capital, deposits, and other inputs partly in response to productivity conditions observed by bank managers but not by the econometrician. Consequently, unobserved productivity is likely to be correlated with observed input choices, and conventional frontier specifications that treat inputs as exogenous may produce biased estimates of production technology and productivity. Tsionas & Mallick (2019) address this problem using the first-order conditions implied by firms' profit-maximization decisions. In our setting, we address the related simultaneity problem by incorporating a control function into a multi-output input distance function.

The fourth issue concerns the operating environment, particularly funding stability, credit risk, and asset quality. Looking beyond banking, Apeti et al. (2024) show that economic and institutional conditions, including trade globalization, factor productivity, and institutional quality, are important determinants of measured public-sector efficiency across 158 countries. Within banking, Ahamed et al. (2021) find that greater financial inclusion is associated with higher bank efficiency, partly through more stable deposit funding. Ahamed & Mallick (2017) show that regulatory forbearance increased the stability of participating banks through lower provisioning requirements on restructured loans, although this effect weakened among banks with greater market power and riskier asset portfolios. At the macroeconomic level, Creel et al. (2023) find that banking fragility, measured by nonperforming loans, reduces economic growth and that credit contributes positively to growth only when banking fragility is relatively low. Collectively, these findings indicate that the economic value of financial intermediation depends not only on the quantity of credit supplied but also on the institutional, funding, and balance-sheet conditions under which it is produced. Accordingly, we include the ratio of loan-loss reserves to gross loans as an environmental variable in the input distance function. This ratio proxies for banks' asset-quality and provisioning conditions, allowing us to examine whether these conditions are associated with the productivity performance of Nigerian banks.

Recent studies further show that changes in banking technology and financial structure do not automatically translate into better intermediation outcomes. Fukuyama et al. (2025) model U.S. bank production using a minimum-distance efficiency framework that treats non-performing loans as an undesirable output and demonstrate how cost skimping can affect long-run performance. Using stochastic frontier models, Al-Azzam et al. (2026) find that global microfinance cost efficiency declined despite evidence of increasing returns to scale. The literature on bank digitalization and fintech similarly emphasizes that technological adoption affects credit outcomes through cost reductions, improved risk management, and reduced information frictions (Wu & Wu, 2025; Feng et al., 2025). These findings reinforce the need to distinguish changes in production technology from changes in banks' relative operating performance.

Against this background, our study makes four contributions. First, it provides new evidence on the dynamics of bank productivity and technical change in Nigeria, a major African banking market that remains underrepresented in the frontier literature. Second, it models banks as multi-output institutions using an input distance function rather than imposing a single-output technology. Third, it incorporates a control function to address the endogeneity arising from banks' input choices. Fourth, it jointly examines productivity change, technical change, scale economies, and input-cost pressures, thereby providing evidence on why consolidation and technological upgrading may not automatically translate into productivity gains. Taken together, the literature leaves limited evidence on how productivity, technical change, scale economies, and input endogeneity interact within a multi-output banking technology in developing economies. We address this gap using bank-level evidence from Nigeria.

# 3 Nigerian Banking Sector

Nigeria's banking industry has undergone substantial reform and structural transformation since the establishment of its first bank in the late 19th century. The Central Bank of Nigeria (CBN), established under the Central Bank Act of 1958 and operational from 1 July 1959, has played a pivotal role in shaping the sector through deregulation, recapitalization, and prudential oversight. Following the return to civilian rule in 1999, Nigeria adopted a universal banking framework in 2001, allowing banks to operate across commercial, investment, and non-banking financial activities (Central Bank of Nigeria, 2000). This framework was repealed in 2010 in favor of a more specialized banking structure (Central Bank of Nigeria, 2010).

The Nigerian banking system comprises commercial, merchant, non-interest (Islamic), microfinance, and development banks. Although microfinance banks are considerably more numerous (Central Bank of Nigeria, 2024, pp. 14–15), commercial, merchant, and non-interest banks—collectively referred to as Deposit Money Banks (DMBs)—remain the principal institutions for financial intermediation. As of 31 December 2024, DMB assets stood at approximately ₦160.6 trillion (Nigeria Deposit Insurance Corporation, 2024, pp. 34–35).

The commercial-banking segment is also highly concentrated. At the end of 2024, the five banks designated by the Central Bank of Nigeria as domestic systemically important banks collectively accounted for 62.81% of industry assets, 61.45% of deposits, and 60.52% of industry credit (Central Bank of Nigeria, 2024, p. 64). In Nigerian financial commentary, these five prominent banking groups are commonly referred to by the informal acronym FUGAZ.[1] Nigerian commercial banks also maintain a substantial international presence. At the end of 2024, the seven banks with international authorization operated 61 offshore affiliates (Central Bank of Nigeria, 2024, p. 14).

The sector experienced a major banking crisis in 2009, rooted in rapid credit expansion, weak corporate governance, and inadequate regulatory oversight during the 2004–2008 period. These vulnerabilities were amplified by the global financial crisis (GFC), which triggered liquidity shortages and a sharp deterioration in asset quality (Sanusi, 2010). The key stages of the crisis and the associated policy responses were as follows.

- Pre-crisis expansion (2004–2008). In 2004, the CBN implemented a consolidation policy that raised the minimum capital requirement from ₦2 billion to ₦25 billion, triggering mergers and

[1] FUGAZ refers to First HoldCo Plc, United Bank for Africa Plc, Guaranty Trust Holding Company Plc, Access Holdings Plc, and Zenith Bank Plc. The term is an informal market designation rather than an official regulatory classification.

acquisitions that reduced the number of banks from 89 to 25. The sector expanded rapidly, with banks engaging heavily in speculative lending, particularly to the stock market. Between 2006 and 2007, equity prices surged, driven largely by margin lending and excessive credit growth.

- Market collapse and liquidity stress (2008–2009). The global financial crisis led to a sharp decline in capital inflows and oil prices, Nigeria's primary revenue source. The stock market lost nearly 70% of its value, exposing banks' vulnerabilities to capital market shocks (Sanusi, 2010). Rising non-performing loans generated widespread liquidity stress, and by mid-2009 several banks were effectively insolvent.

- Regulatory intervention (2009). In August 2009, the CBN injected ₦620 billion to stabilize eight distressed banks, replaced senior management teams, and introduced reforms aimed at strengthening supervision, corporate governance, and risk management (Sanusi, 2010).

In 2010, the Asset Management Corporation of Nigeria (AMCON) was established to absorb non-performing loans and impaired assets and further stabilize the banking system (Sanusi, 2010). Since then, the banking sector, dominated by commercial banks, has remained a key pillar of the Nigerian economy. In parallel, banks have accelerated the adoption of digital technologies, including mobile and USSD banking[2], agent banking, AI-driven customer service tools, and the introduction of the eNaira, reflecting a transition toward a more technology-driven and service-oriented financial system. Given the sector's scale and systemic importance, understanding productivity dynamics in Nigerian banking is essential for assessing financial performance and supporting sustainable economic development.

# 4 Bank Production Model

We model banks as production units engaged in financial intermediation, generating multiple outputs, such as loans and investments, using inputs including deposits & borrowings, capital, and labor. Since banking outputs are non-storable services and are treated as exogenous, we assume that banks choose inputs to minimize variable costs. Furthermore, because input price data are unavailable, we estimate production technology using an IDF derived from a multiple-output, multiple-input transformation function. The IDF is dual to the cost function, meaning that cost-function elasticities can be recovered directly from the IDF.

## 4.1 Transformation Technology

We specify the transformation function, augmented to include neutral productivity, as

$$F(Y_{it}, X_{it}e^{\omega_{it}}, Z_{it}, t) = 1, \tag{1}$$

[2] Unstructured Supplementary Service Data (USSD) is a mobile telecommunications protocol that enables transactions via shortcodes without requiring internet access.

where $Y_{it}$ denotes the vector of outputs, $X_{it}$ is the vector of variable inputs, $Z_{it}$ denotes an environmental variable such as non-performing loans,[3] $t$ is a time trend that captures technical change, and $\omega_{it}$ is a neutral productivity term that scales all inputs symmetrically. The following assumptions are made on the technology in (1).

*Assumption 1.* Banks produce $M$ outputs using $J$ variable inputs. Outputs, such as net loans and investments,[4] are exogenous to the bank. This treatment follows the theory of financial intermediation, in which loan supply is influenced by external factors, such as the business cycle and monetary policy, rather than by contemporaneous input choices.

*Assumption 2.* Productivity $\omega_{it}$ is observed by banks at the time of input decisions but is unobserved by the econometrician. It captures persistent productivity differences across banks and evolves according to a first-order Markov process, as is standard in the production function estimation literature (Olley & Pakes, 1996; Levinsohn & Petrin, 2003; Ackerberg et al., 2015):

$$\omega_{it} = \rho_1 \omega_{it-1} + \delta_{GFC}\, GFC_{it} + \delta_G \ln G_{it-1} + \xi_{it}, \tag{2}$$

where $\xi_{it}$ is a productivity innovation that is mean-independent of the bank's information set at $t-1$, i.e., $E[\xi_{it} \mid \mathcal{I}_{it-1}] = 0$, $GFC_{it}$ is an indicator for the global financial crisis period, and $\ln G_{it-1}$ is the lagged logarithm of total assets. In addition, $\ln G_{it-1}$ is included as a productivity shifter to capture size, balance-sheet capacity, and scale-related heterogeneity across banks. The term $GFC_{it}$ is included as a productivity shifter to capture the effects of the global financial crisis, which propagated to the Nigerian banking sector between 2008 and 2010.[5]

*Assumption 3.* Banks minimize costs by choosing inputs to produce a given vector of exogenous outputs in period $t$.

*Assumption 4.* The transformation function $F(\cdot)$ is homogeneous of degree one in the variable inputs. This is an identifying assumption on the technology (Kumbhakar, 2013).

Using the homogeneity assumption in Assumption 4 and choosing $X_{1,it}e^{\omega_{it}}$ as the numeraire, we rewrite (1) as

$$F\left(Y_{it},\ 1,\ \frac{X_{-1,it}}{X_{1,it}},\ Z_{it},\ t\right) = \frac{1}{X_{1,it}\, e^{\omega_{it}}}, \tag{3}$$

where $X_{-1,it}$ includes all variable inputs except $X_{1,it}$. Note that $e^{\omega_{it}}$ cancels from the input ratios on the left-hand side because productivity is neutral. Taking logarithms yields the standard IDF:

$$\ln X_{1,it} = \ln D(Y_{it}, 1, \tilde{X}_{it}, Z_{it}, t) - \omega_{it}, \tag{4}$$

[3]The ratio of reserve for loan losses to gross loans is used as a proxy for non-performing loans in our model.

[4]We use the term "investment" instead of "securities" because this is the terminology used in the balance sheets of Nigerian banks.

[5]The AR(1) process is specified without an intercept, as any constant term would be indistinguishable from the production-function constant $\beta_0$ (Ackerberg et al., 2015).

where $\ln D(Y_{it}, 1, \tilde{X}_{it}, Z_{it}, t) \equiv -\ln F(Y_{it}, 1, \tilde{X}_{it}, Z_{it}, t)$ and $\tilde{X}_{it} = \left(\frac{X_{2,it}}{X_{1,it}}, \ldots, \frac{X_{J,it}}{X_{1,it}}\right)$.

Because banks observe productivity when selecting inputs, $\omega_{it}$ can be correlated with the input ratios $\tilde{X}_{it}$ and therefore cannot be treated as a regression error. Using the AR(1) structure in (2), equation (4) can be rewritten as

$$\ln X_{1,it} = \ln D(Y_{it}, 1, \tilde{X}_{it}, Z_{it}, t) - g(\omega_{it-1}, GFC_{it}, \ln G_{it-1}) - \xi_{it}, \tag{5}$$

where $g(\omega_{it-1}, GFC_{it}, \ln G_{it-1}) = \mathbb{E}[\omega_{it} \mid \mathcal{I}_{it-1}, GFC_{it}] = \rho_1 \omega_{it-1} + \delta_{GFC}\, GFC_{it} + \delta_G \, \ln G_{it-1}$.

We employ a translog functional form for $\ln D(\cdot)$ because of its flexibility in modeling production technology. Unlike the Cobb–Douglas specification, the translog accommodates varying elasticities of substitution, returns to scale, and cost elasticities.[6] Denoting $\mathcal{Q}_{it} = (\ln Y'_{it},\, \ln \tilde{X}'_{it},\, Z'_{it},\, t)'$, which collects all right-hand-side variables (recall that $Z_{it}$ is a ratio and enters in levels rather than logs), we write the translog IDF as

$$\ln X_{1,it} = \beta_0 + \sum_{j=1}^{N_{\mathcal{Q}}} \beta_{\mathcal{Q},j}\, \mathcal{Q}_{j,it} + \frac{1}{2} \sum_{j=1}^{N_{\mathcal{Q}}} \sum_{k=1}^{N_{\mathcal{Q}}} \beta_{\mathcal{Q},jk}\, \mathcal{Q}_{j,it}\, \mathcal{Q}_{k,it} - \rho_1\, \omega_{it-1} - \delta_{GFC}\, GFC_{it} - \delta_G \, \ln G_{it-1} - \xi_{it}, \tag{6}$$

where $N_{\mathcal{Q}}$ denotes the dimension of the vector $\mathcal{Q}_{it}$ and the symmetry restriction $\beta_{\mathcal{Q},jk} = \beta_{\mathcal{Q},kj}$ is imposed.

In the estimated specification, the environmental variable $Z_{it}$ enters only through its own linear and quadratic terms and is not interacted with $Y_{it}$, $\tilde{X}_{it}$, or $t$; hence the 22 frontier coefficients consist of the 20 output, normalized-input, and time terms listed in $\mathcal{B}_{it}$ below plus the two own terms in $Z_{it}$.

All variables in $\mathcal{Q}_{it}$ are exogenous or predetermined except for $Z_{it}$, which can be endogenous: outputs are exogenous by Assumption 1, and the first-order conditions of cost minimization ensure that input ratios are also exogenous (Wang et al., 2025). This guarantees identification of (6) (see Wang et al. (2025) for details). The only issue is the presence of the unobserved $\omega_{it-1}$, which must be substituted out in terms of observables before the equation can be estimated.

## 4.2 Estimation

Our focus is on estimating the IDF in equation (6), from which we recover $\omega_{it}$ and various economic measures. From equation (4), lagged productivity can be expressed as $\omega_{it-1} = \ln D(\mathcal{Q}_{it-1}) - \ln X_{1,it-1}$.

[6] We use the Cobb–Douglas specification as a robustness check to assess the sensitivity of our results to a more restrictive functional form.

Substituting this into (6) and collecting terms yields

$$\begin{aligned}
\ln X_{1,it} &= \beta_0 + \sum_{j=1}^{N_{\mathcal{Q}}} \beta_{\mathcal{Q},j}\, \mathcal{Q}_{j,it} + \tfrac{1}{2} \sum_{j=1}^{N_{\mathcal{Q}}} \sum_{k=1}^{N_{\mathcal{Q}}} \beta_{\mathcal{Q},jk}\, \mathcal{Q}_{j,it}\, \mathcal{Q}_{k,it} \\
&\quad - \rho_1\, \omega_{it-1} - \delta_{GFC}\, GFC_{it} - \delta_G\, \ln G_{it-1} - \xi_{it} \\
&= \beta_0 + \sum_{j=1}^{N_{\mathcal{Q}}} \beta_{\mathcal{Q},j}\, \mathcal{Q}_{j,it} + \tfrac{1}{2} \sum_{j=1}^{N_{\mathcal{Q}}} \sum_{k=1}^{N_{\mathcal{Q}}} \beta_{\mathcal{Q},jk}\, \mathcal{Q}_{j,it}\, \mathcal{Q}_{k,it} \\
&\quad - \rho_1 \Big( - \ln X_{1,it-1} + \beta_0 + \sum_{j=1}^{N_{\mathcal{Q}}} \beta_{\mathcal{Q},j}\, \mathcal{Q}_{j,it-1} \\
&\qquad + \frac{1}{2} \sum_{j=1}^{N_{\mathcal{Q}}} \sum_{k=1}^{N_{\mathcal{Q}}} \beta_{\mathcal{Q},jk}\, \mathcal{Q}_{j,it-1}\, \mathcal{Q}_{k,it-1} \Big) \\
&\quad - \delta_{GFC}\, GFC_{it} - \delta_G\, \ln G_{it-1} - \xi_{it} \\
&= \beta_0(1-\rho_1) + \rho_1 \ln X_{1,it-1} + \sum_{j=1}^{N_{\mathcal{Q}}} \beta_{\mathcal{Q},j} \big[ \mathcal{Q}_{j,it} - \rho_1\, \mathcal{Q}_{j,it-1} \big] \\
&\quad + \frac{1}{2} \sum_{j=1}^{N_{\mathcal{Q}}} \sum_{k=1}^{N_{\mathcal{Q}}} \beta_{\mathcal{Q},jk} \Big( \mathcal{Q}_{j,it}\, \mathcal{Q}_{k,it} - \rho_1\, \mathcal{Q}_{j,it-1}\, \mathcal{Q}_{k,it-1} \Big) \\
&\quad - \delta_{GFC}\, GFC_{it} - \delta_G\, \ln G_{it-1} - \xi_{it},
\end{aligned} \tag{7}$$

where $\omega_{it-1} = \ln D(\mathcal{Q}_{it-1}) - \ln X_{1,it-1}$.

Equation (7) is expressed entirely in terms of observables and the productivity innovation $\xi_{it}$. Since outputs are exogenous (Assumption 1) and input ratios are exogenous under cost minimization (Wang et al. (2025)), the following orthogonality conditions hold:

$$\mathbb{E}[\mathcal{H}_{it} \cdot \xi_{it}] = 0, \tag{8}$$

where $\mathcal{H}_{it}$ is the instrument vector defined below. The equation can therefore be estimated consistently by nonlinear least squares (NLS) or, to improve efficiency, by the generalized method of moments (GMM).

### 4.3 Instruments

The vector $\mathcal{B}_{it}$ comprises five base variables — two output measures, two input ratios, and a time trend — along with their associated square and interaction terms, i.e.,

$$\mathcal{B}_{it} = \Big(\ln \tilde{X}_{2,it},\ \ln \tilde{X}_{3,it},\ \ln Y_{1,it},\ \ln Y_{2,it},\ t,\ (\ln \tilde{X}_{2,it})^2,\ (\ln \tilde{X}_{3,it})^2,\ (\ln Y_{1,it})^2,\ (\ln Y_{2,it})^2,\ t^2,$$
$$\ln \tilde{X}_{2,it} \ln \tilde{X}_{3,it},\ \ln \tilde{X}_{2,it} \ln Y_{1,it},\ \ln \tilde{X}_{2,it} \ln Y_{2,it},\ \ln \tilde{X}_{2,it}\, t,\ \ln \tilde{X}_{3,it} \ln Y_{1,it},\ \ln \tilde{X}_{3,it} \ln Y_{2,it},$$
$$\ln \tilde{X}_{3,it}\, t,\ \ln Y_{1,it} \ln Y_{2,it},\ \ln Y_{1,it}\, t,\ \ln Y_{2,it}\, t\Big)',$$

giving 20 distinct elements (5 linear, 5 squared, and 10 cross-products). Because $Z_{it}$ is constructed as the ratio of reserve for loan losses to gross loans, and reserve for loan losses might be endogenous, we use its lag as an instrument so that it does not correlate with the productivity innovation $\xi_{it}$.

The lagged vector $\mathcal{B}_{it-1}^{-t}$ contains all elements of $\mathcal{B}_{it-1}$ with the exception of $t-1$ and $(t-1)^2$. These are excluded to avoid perfect collinearity with the contemporaneous time trend terms ($t$ and $t^2$) already present in the instrument set. Additionally, we include $Z_{it-1}$, $Z_{it-1}^2$, and $\ln G_{it-1}$ in $\mathcal{B}_{it-1}^{-t}$, resulting in 21 lagged elements. The instrument vector is defined as:

$$\mathcal{H}_{it} = \Big(1,\ \ln X_{1,it-1},\ \mathcal{B}_{it}',\ \mathcal{B}_{it-1}^{-t\,\prime}\Big)', \tag{9}$$

yielding $1+1+20+21 = 43$ moment conditions. These conditions identify 26 parameters—comprising $\beta_0$, 22 translog coefficients, $\rho_1$, $\delta_{\mathrm{GFC}}$, and $\delta_{\mathrm{G}}$—leaving 17 overidentifying restrictions testable via Hansen's $J$-statistic.

### 4.4 Derived Measures

Because the IDF is dual to the cost function (Färe & Primont, 1995), it directly delivers cost elasticities and the rate of technical change (TC):

$$E_{CY_m} = \frac{\partial \ln C}{\partial \ln Y_m} = \frac{\partial \ln X_1}{\partial \ln Y_m}, \quad m = 1, \ldots, M, \qquad \frac{\partial \ln C}{\partial t} = \frac{\partial \ln X_1}{\partial t}. \tag{10}$$

These cost elasticities characterize returns to scale (RTS), viz., $RTS = \{\sum_m E_{CY_m}\}^{-1}$. Bank-level productivity is recovered residually as

$$\hat{\omega}_{it} = \ln \hat{D}(\cdot) - \ln X_{1,it}. \tag{11}$$

### 4.5 Banking Approach and Variable Classification

The banking literature offers two main microeconomically grounded frameworks for modeling banks: the production approach and the intermediation approach. The production approach, introduced by Benston

(1965), views banks as producers of financial services, where inputs include labor, materials, and physical facilities and outputs represent the services provided to customers. In contrast, the intermediation approach conceptualizes banks as financial intermediaries that mobilize deposits and other liabilities from surplus units and channel them into loans and investments. A central debate concerns the classification of deposits: whether they function as inputs, reflecting the cost of mobilizing funds, or as outputs, representing the services provided to depositors (Berger & Humphrey, 1997; Das & Kumbhakar, 2012).

Berger & Humphrey (1997) argue that the production approach is most appropriate for branch-level analysis, whereas the intermediation approach is better suited for system-wide studies, primarily because many balance-sheet items, particularly deposits, have dual roles (Berger et al., 1993). Deposits function as inputs when used as loanable funds and as outputs when associated with liquidity provision and payment services.

This study adopts the intermediation approach, treating deposits & borrowings as inputs because the objective is to evaluate productivity at the banking-system level. In the baseline specification, inputs include capital, labor, and deposits & borrowings, while outputs consist of loans and investments. The ratio of reserve for loan losses to gross loans is treated as an environmental variable to control for credit-quality deterioration, that is, the presence of impaired output that affects the production process. The ratio form is used rather than the level of reserves because it nets out differences in bank size that are inherent in the level data, thereby isolating the intensity of credit impairment relative to each bank's lending activity.

## 5 Data

The empirical analysis employs an unbalanced panel dataset constructed from the annual reports of Nigerian commercial banks covering the period 2000–2023. The panel is unbalanced because banks differ in the years for which financial information is available. Although the primary data source is the published annual reports of the banks, access to these reports was facilitated through the Bloomberg Financial Report on Nigerian banks.[7] The cleaned panel contains 273 bank-year observations. For the baseline estimation, monetary variables are first deflated, and observations in the top and bottom 1% tails of the logged variables entering the model are trimmed before estimation. The 1% trim follows the common practice in the literature for handling outliers and was fixed as the baseline before the alternative treatments in Section 6.6.8 were examined. All monetary variables used in the analysis are deflated to real values using the Nigerian GDP deflator from the International Monetary Fund World Economic Outlook Database (International Monetary Fund, 2025).[8] Missing observations for specific

[7] The dataset was accessed via the Bloomberg Terminal at the State University of New York at Binghamton.

[8] The deflator is the WEO subject-code series `NGDP_D`, "Gross domestic product, deflator," for Nigeria. The extracted series is an index with 2010 = 100, and monetary variables are converted to constant 2010 naira as nominal values divided by the deflator index over 100. We use the GDP deflator index, not the annual percentage inflation rate, for deflating level variables.

years were supplemented with information obtained directly from the banks' financial statements.

We initially collected data on 19 banks: one foreign bank, one Islamic bank, and 17 commercial banks. We exclude the foreign and Islamic banks because their operations differ fundamentally from those of conventional commercial banking, leaving a study panel of 17 commercial banks. After trimming, lag construction, and complete-case filtering, the baseline dynamic GMM estimation uses 230 observations from 15 banks. Access Bank and Titan Trust remain in the study panel but contribute no usable baseline GMM rows because their short series and tail observations do not leave the required consecutive complete records. Of these 17 banks, 15 were still operating at the end of the sample period, while the other two had been acquired by other banks during the sample period but were retained because they reported data for several years prior to acquisition. In our final sample, 10 are internationally licensed banks (ILBs) and 7 are nationally licensed banks (NLBs).[9] We further classify banks by size into FUGAZ and non-FUGAZ categories based on their total assets and market capitalization. These five institutions account for the largest share of assets and market capitalization in the Nigerian banking sector, while the remaining 12 banks constitute the non-FUGAZ group.

We excluded recent entrants to commercial banking (earliest operations in 2016–2023) due to a lack of sufficient operating histories. Many have fewer than three years of observations, which is inadequate for our GMM estimator's dynamic panel moment conditions. Foreign-owned subsidiaries that focus on corporate and trade finance rather than full-service commercial banking and report consolidated financial statements at the parent-group level were also excluded. These excluded institutions represent a small share of commercial banking assets, whereas our 17 retained banks include all systemically important banks in Nigeria and represent the overwhelming majority of total industry assets.

The dataset includes the following variables: labor (number of staff), capital (net fixed assets), deposits and borrowings, net loans, long-term and short-term investments, total assets, reserve for loan losses, and the ratio of reserve for loan losses to gross loans (RFLL). Net loans are defined as total loans minus reserves for loan losses. Total investments (denoted as "investment") represent the sum of long-term and short-term investments. Although the initial objective was to distinguish multiple categories of loans and investments, inconsistencies in variable definitions across bank reports and missing data necessitated the use of common variables across banks. Observations missing critical variables, deposits & borrowings, investments, or net loans were excluded, yielding the cleaned panel described above.

## 5.1 Summary Statistics

Table 1 presents the summary statistics for the main inputs, outputs, and additional banking variables used in the analysis. All monetary variables are measured in millions of constant 2010 naira (₦), while labor is measured as the total number of staff. Several clear patterns emerge from the table.

First, the Nigerian banking sector is characterized by substantial heterogeneity in bank size and

[9]Only commercial banks are included; microfinance, development, and merchant banks are excluded from the analysis.

balance-sheet composition. Deposits & borrowings, which constitute the main funding input under the intermediation approach, have a mean of ₦900.16 billion, a median of ₦528.25 billion, and a standard deviation of ₦1.07 trillion. The large gap between the mean and median, together with the very high dispersion, indicates a strongly right-skewed distribution in which a small number of very large banks account for a disproportionate share of sectoral funding.

Second, a similar pattern is observed for both outputs and other balance-sheet variables. Mean net loans amount to ₦415.11 billion, compared with a median of ₦248.34 billion, while mean investment is ₦268.91 billion relative to a median of ₦137.31 billion. Total assets also exhibit wide dispersion, with a mean of ₦1.14 trillion, a median of ₦747.44 billion, and a standard deviation of ₦1.28 trillion. These differences confirm that the sample includes banks operating at very different scales, ranging from relatively small domestic institutions to much larger banks with broader balance-sheet capacity.

Third, the input variables labor and capital also vary substantially across banks. The average number of staff is 4,152, while the median is 3,056, suggesting that employment size is also skewed toward a few large institutions. Net fixed assets (capital) have a mean of ₦30.20 billion and a standard deviation of ₦25.34 billion, again pointing to large variation in physical capacity and branch infrastructure across banks. This is consistent with the institutional structure of the Nigerian banking industry, where larger banks tend to maintain wider branch networks, stronger capital bases, and greater operational reach.

Finally, reserve for loan losses has a mean of ₦23.48 billion and a median of ₦13.75 billion, with a standard deviation of ₦29.92 billion. This suggests substantial cross-bank differences in credit risk exposure and loan portfolio quality. The fact that the mean exceeds the median by a wide margin implies that a subset of banks carries much larger provisions than the rest of the sample.

Overall, the summary statistics indicate that the Nigerian banking sector is highly concentrated and unevenly distributed in terms of funding, asset holdings, lending activity, investment, and risk exposure. The persistent pattern in which means exceed medians across most financial variables suggests right-skewed distributions dominated by a small number of large institutions. This degree of heterogeneity provides strong support for the use of a flexible multi-output framework, since banks in the sample differ markedly in scale, portfolio structure, and productive capacity.

Table 1: Summary Statistics

| | 1st Qu. | Median | Mean | SD | 3rd Qu. |
|---|---|---|---|---|---|
| Deposits & borrowings (X1) | 234,280 | 528,252 | 900,160 | 1,068,912 | 1,129,460 |
| Labor (X2) | 1,832 | 3,056 | 4,152 | 3,154 | 6,681 |
| Capital (Net Fixed Assets, X3) | 10,591 | 20,439 | 30,198 | 25,339 | 47,093 |
| Loans (Y1) | 98,869 | 248,340 | 415,110 | 456,691 | 593,464 |
| Investment (Y2) | 54,155 | 137,314 | 268,905 | 360,516 | 307,351 |
| Total Assets (G) | 281,917 | 747,440 | 1,141,589 | 1,276,560 | 1,468,955 |
| Reserve for Loan Losses | 4,981 | 13,751 | 23,478 | 29,924 | 32,158 |
| RFLL Ratio (Z) | 0.0248 | 0.0417 | 0.0756 | 0.0875 | 0.0888 |

*Notes:* All monetary variables are measured in millions of constant 2010 naira, while labor is measured as the number of staff. The RFLL ratio is the ratio of reserve for loan losses to gross loans. All variables are bank-year observations for 17 commercial banks over 2000–2023.

# 6 Results

In this section, we estimate the input distance function in equation (7) using nonlinear generalized method of moments (GMM). Our primary objects of interest are estimates of change in productivity, output cost elasticities, returns to scale, and technical change. We subject the baseline results to a battery of robustness checks. Our estimation imposes a common production technology across all banks in the sample. This assumption is justified on institutional grounds: all banks in the Nigerian banking sector operate under the same regulatory frameworks and are conventional commercial banks. Under the common technology, differences in estimated productivity across banks reflect variation in efficiency rather than differences in the underlying production function. Although our sample spans 2000 to 2023, the reported productivity-change, RTS, and technical-change summaries begin in 2002 because of the lag structure and dynamic differences in the model. Some observation-level elasticity and RFLL plots include 2001 when the underlying estimated quantity is available. Unless otherwise stated, percentages in the text are computed from the unrounded estimates after multiplying by 100 and are reported to one decimal place; exact half-way cases are rounded away from zero, while tables report values at the precision displayed.

## 6.1 Elasticities

Table 2 and Figure 2 report the estimated output cost elasticities for loans, $\epsilon_1$, and investment, $\epsilon_2$. In the IDF, input-ratio derivatives describe how required deposits and borrowings vary with the input mix, while positive output elasticities indicate that higher output requires proportionally greater input usage.

The median elasticity of loans is 0.4897, about 2.4 times the median elasticity of investment of 0.2024. The same ordering holds at the mean, where the loan elasticity is 0.5010 compared with 0.1870 for investment. This implies that loans are the dominant cost driver in Nigerian banking: a one percent increase in loan output requires approximately a 0.49 to 0.50 percent proportional increase in inputs, whereas the same percentage increase in investment requires about a 0.19 to 0.20 percent increase. The result is consistent with the intermediation view of banking, in which lending, with its associated credit screening, monitoring, and provisioning costs, constitutes the core and most resource-intensive activity.

The subgroup decomposition in Panels A and B of Figure 2 reveals meaningful but more nuanced heterogeneity. Along the size dimension, Non-FUGAZ banks have a median loan elasticity of 0.507, compared with 0.472 for FUGAZ banks, indicating that lending is cost-intensive throughout the sector and slightly more input-demanding for smaller banks. Investment elasticities are much closer across size groups, with medians of 0.203 for FUGAZ and 0.198 for Non-FUGAZ banks. Along the license dimension, ILBs display a higher median loan elasticity of 0.501 compared with 0.467 for NLBs, while NLBs have the higher median investment elasticity, 0.217 compared with 0.187 for ILBs.

Panels C through F of Figure 2 show that the gap between loan and investment elasticities narrows over time but does not disappear. The average loan elasticity is around 0.60 in the early 2000s and declines to around 0.40–0.45 in the later years, while the investment elasticity rises from roughly 0.15 in the early years to about 0.22–0.25 in much of the post-2010 period. This pattern is consistent with a gradual reduction in the input intensity of lending, plausibly reflecting credit-scoring, electronic origination, and monitoring technologies, alongside an increase in the relative complexity of investment activities. Even in the later period, however, lending remains the larger cost component.

### 6.1.1 Input Elasticities

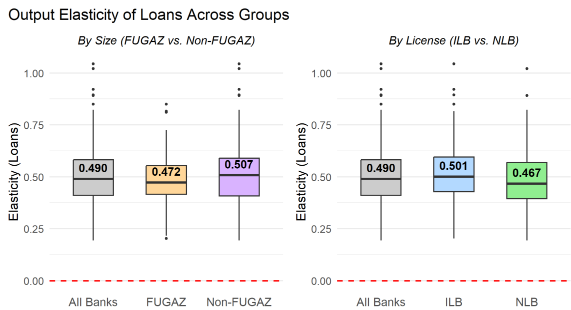
From the IDF specification, we define $m_{2,it} = \frac{\partial \ln X_{1,it}}{\partial \ln \tilde{X}_{2,it}}$ and $m_{3,it} = \frac{\partial \ln X_{1,it}}{\partial \ln \tilde{X}_{3,it}}$, which represent the percentage change in deposits and borrowings, $X_1$, resulting from a one percent change in the input ratios $X_2/X_1$ (labor to deposits and borrowings) and $X_3/X_1$ (capital to deposits and borrowings), respectively. Table 2 reports that $m_{2,it}$ has a mean of $-0.2505$, while $m_{3,it}$ has a mean of $-0.1957$. Both mean estimates are negative, indicating that an increase in either input ratio is associated with a reduction in the use of deposits and borrowings.

Since interpretations of $m_{2,it}$ and $m_{3,it}$ are not straightforward, we use regular input elasticities $E_{2,it}$ and $E_{3,it}$ defined as:

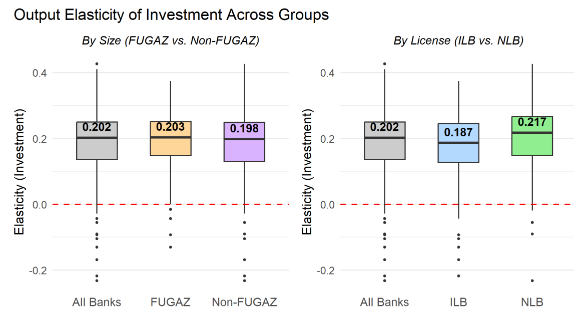
$$E_{2,it} \equiv \frac{\partial \ln X_{2,it}}{\partial \ln X_{1,it}}, \quad E_{3,it} \equiv \frac{\partial \ln X_{3,it}}{\partial \ln X_{1,it}}.$$

These elasticities are obtained using the relationship $E_{j,it} = (1 + m_{j,it})/m_{j,it}$, $j = 2, 3$. $E_{2,it}$ and $E_{3,it}$ capture the percentage change in labor, $X_2$, and capital, $X_3$, in response to a one percent change in deposits and borrowings, $X_1$. A negative (positive) value of $E_{j,it}$ indicates that $X_{j,it}$ and $X_{1,it}$ are substitutes (complements). Evaluated at the mean values of $m_2$ and $m_3$, the implied input elasticities are approximately $-2.99$ for $E_2$ and $-4.11$ for $E_3$. Both negative values indicate that labor and capital are substitutes for deposits and borrowings.

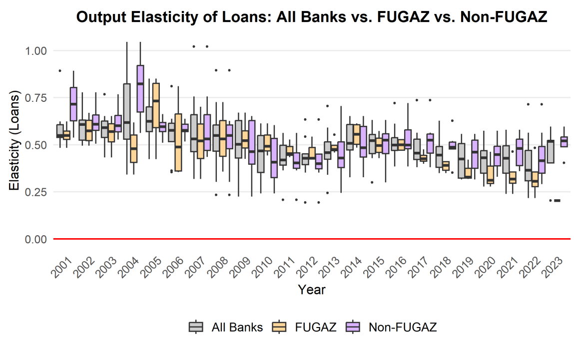
The substitutability between deposits & borrowings and labor, reflected in $E_{2,it} \approx -2.99$, is consistent with the ongoing shift toward digital banking channels, automated teller machines, mobile platforms, and back-office automation that allow banks to mobilize and service deposits with fewer staff-intensive processes. The substitutability between deposits and capital, with $E_{3,it} \approx -4.11$, is stronger, suggesting that deposit growth is increasingly supported by scalable infrastructure rather than proportional expansion of physical capital. These findings are consistent with the broader dynamics of input utilization in banking documented by Wang et al. (2025). The final row in Table 2, labeled "technical change," is explained in the technical change subsection of the results.



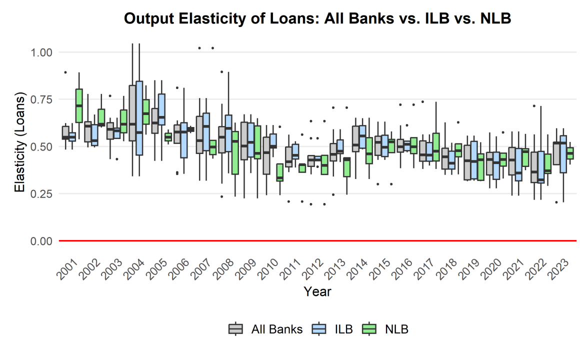

**Panel A:** Output Elasticity of Loans



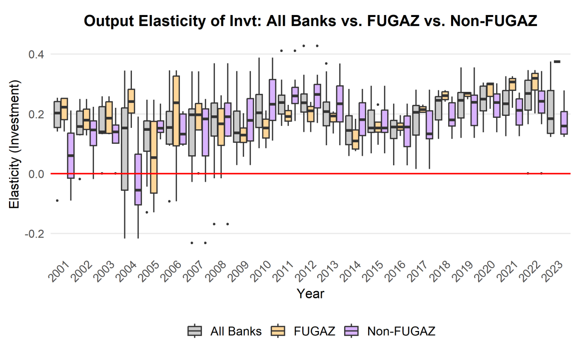

**Panel B:** Output Elasticity of Investment



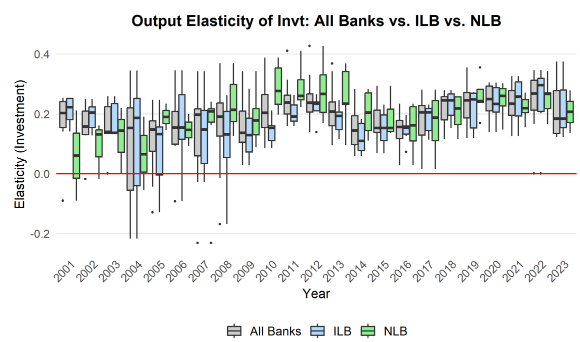

**Panel C:** Loans: All Banks vs. FUGAZ vs. Non-FUGAZ




**Panel D:** Loans: All Banks vs. ILB vs. NLB




**Panel E:** Investment: All Banks vs. FUGAZ vs. Non-FUGAZ




**Panel F:** Investment: All Banks vs. ILB vs. NLB

Figure 2: Output Elasticity of Loans and Investments

Table 2: Summary of Elasticities (All Banks)

| | Q1 | Median | Mean | Q3 |
|---|---|---|---|---|
| $m_2$ (Labor/Deposits & Borrowing) | $-0.3558$ | $-0.2348$ | $-0.2505$ | $-0.1370$ |
| $m_3$ (Capital/Deposits & Borrowing) | $-0.2883$ | $-0.2019$ | $-0.1957$ | $-0.1215$ |
| $\epsilon_1$ (Loans) | 0.4105 | 0.4897 | 0.5010 | 0.5819 |
| $\epsilon_2$ (Investment) | 0.1355 | 0.2024 | 0.1870 | 0.2501 |
| TC (Technical Change) | $-0.0485$ | $-0.0354$ | $-0.0351$ | $-0.0223$ |

## 6.2 Productivity

We measure productivity using the estimated $\omega_{it}$ and productivity changes from $\Delta\omega_{it} = \omega_{it} - \omega_{it-1}$. We examine the trajectory of productivity change across time, bank groups, and key regulatory episodes in turn.

### 6.2.1 Productivity Dynamics: 2002 to 2023

Between 2002 and 2023, Nigerian banks exhibited an average annual productivity decline of 3.5%, with a median decline of 3.1% (see Table 3 and Figure 3). The negative mean and median confirm that the typical bank in the pooled sample experienced productivity erosion over the sample period. The subgroup patterns below show steeper losses among the largest and internationally licensed banks, while the sample-restriction checks in Section 6.6.7 show that the decline in mean productivity is not confined to those groups.

The decline is not uniform across bank groups. Along the size dimension, FUGAZ banks recorded a mean (median) productivity change of −5.6% (−5.4%), compared with −2.3% (−0.8%) for Non-FUGAZ banks. Along the license dimension, ILBs recorded a mean (median) productivity change of −5.0% (−3.8%), while NLBs recorded −1.5% (−0.8%). These patterns indicate that the largest and internationally licensed banks experienced larger productivity losses, even though the scale estimates below show that banks operated under increasing returns to scale.

### 6.2.2 Early Period: 2002 to 2005

Panel A of Figure 3 shows that productivity fluctuated during the early part of the sample. Mean productivity change was 2.1% in 2002, −4.6% in 2003, −3.9% in 2004, and 4.0% in 2005. The average productivity change across 2002–2005 was therefore close to zero at −0.6%, indicating that the early reform period was volatile rather than uniformly productivity enhancing.

Subgroup patterns reveal that NLB and Non-FUGAZ banks performed relatively better in this early period. NLBs recorded an average productivity change of 3.2%, compared with −2.2% for ILBs. Similarly, Non-FUGAZ banks recorded an average productivity change of 5.5%, while FUGAZ banks declined by 5.9% on average. The relatively stronger performance of Non-FUGAZ banks and NLBs suggests that smaller and domestically focused banks were able to realize short-run gains during parts of the run-up to consolidation, while larger and internationally licensed banks absorbed higher transition and compliance costs.

### 6.2.3 First Recapitalization: 2004 to 2006

In 2004, the CBN initiated a major recapitalization policy by raising the minimum capital requirement for banks from ₦2 billion to ₦25 billion, triggering a wave of mergers and acquisitions that reduced the number of banks from 89 to 25. The productivity response around this reform was sharply mixed.

In 2005, mean productivity rose by 4.0% overall, with NLBs increasing by 12.4% and ILBs by 1.2%. The gain was concentrated among Non-FUGAZ banks, which recorded a productivity increase of 13.3%, while FUGAZ banks declined by 5.3%.

The positive effect in 2005 was reversed in 2006, when mean productivity fell by 27.3% for all banks. This contraction likely reflects the adjustment costs associated with mergers, operational integration, balance-sheet restructuring, and post-consolidation organizational disruptions. The decline was more severe for ILBs at −32.6% than for NLBs at −11.3%. Along the size dimension, FUGAZ banks declined by 20.6%, while Non-FUGAZ banks declined by 33.9%, indicating that the immediate post-consolidation adjustment burden was broad and not confined to the largest banks.

### 6.2.4 Global Financial Crisis: 2008 to 2010

The global financial crisis period shows a visible but uneven productivity response. Mean productivity declined by 6.7% in 2008 and 1.1% in 2009, before increasing by 2.7% in 2010. The average productivity change over 2008–2010 was −1.7%, suggesting that the crisis period generated moderate productivity pressure on average, although the effect was smaller than the 2006 post-consolidation contraction and was masked by considerable cross-sectional heterogeneity.

The subgroup patterns suggest that FUGAZ and ILBs were relatively more resilient during the crisis period. Over 2008–2010, FUGAZ banks recorded an average productivity change of −0.2%, compared with −2.4% for Non-FUGAZ banks. Along the license dimension, ILBs declined by 0.6% on average, while NLBs declined by 3.1%. This resilience likely reflects the stronger capital buffers, more diversified funding sources, and greater crisis-absorption capacity of larger and internationally licensed banks.

To formally test for a crisis effect, we include a dummy variable for the period 2008 to 2010 in the productivity process. The estimated coefficient $\hat{\delta}_{\mathrm{GFC}}$ is 0.0373 with a $p$-value of 0.506 based on standard errors clustered at the bank level, so the crisis dummy is not statistically significant in the baseline specification. The sign is positive, but the imprecision indicates that the crisis-period shift in productivity levels is not separately identified from the AR(1) dynamics and the lagged total-assets term with enough precision to support a strong direct-crisis interpretation.

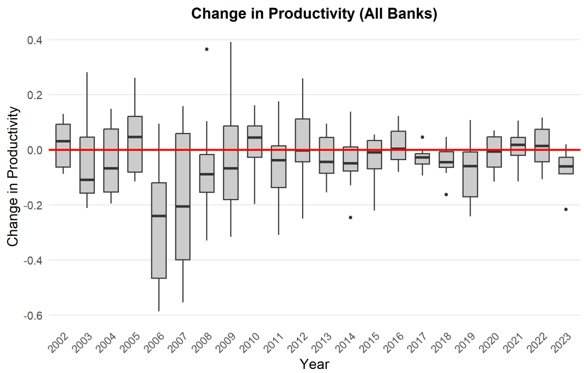
### 6.2.5 Post-Recapitalization Period: 2013 to 2023

After 2012, productivity dynamics stabilized at a persistently negative but less extreme level than in the immediate post-2004 consolidation years. The mean productivity decline over 2013–2023 was 2.7% overall, with ILBs declining by 2.9% and NLBs by 2.5%. FUGAZ banks experienced the largest average decline at 4.3%, compared with 1.8% for Non-FUGAZ banks. Thus, the post-2012 period did not generate sustained productivity gains, but the decline was more moderate than in 2006–2007.

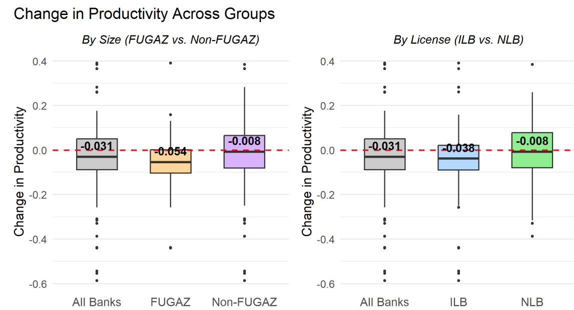
Year-on-year volatility remained visible. Productivity increased in 2016, 2021, and 2022, but these gains did not overturn the negative average for the period. In 2023, average bank-level productivity declined by 7.4%, with the contraction particularly acute for NLBs (−13.8%),[10] followed by FUGAZ banks (−8.7%), Non-FUGAZ banks (−7.1%), and ILBs (−3.1%). The 2023 decline is therefore an important late-sample deterioration, although it is smaller than the severe adjustment losses observed in 2006 and 2007.

The 2023 productivity contraction appears to coincide with a cluster of macroeconomic and regulatory shocks that placed substantial pressure on banks' operating environment. The naira redesign and associated cash scarcity disrupted payment activity, increased transaction costs, and forced a rapid shift toward electronic channels. At the same time, monetary policy tightening, higher cash reserve requirements, exchange-rate depreciation following foreign-exchange market reforms, and the removal of the fuel subsidy increased funding costs, inflationary pressures, and borrower repayment risk. These shocks are consistent with the particularly severe decline observed among nationally licensed banks, which tend to have thinner buffers and more limited capacity to absorb system-wide operational and macroeconomic stress.

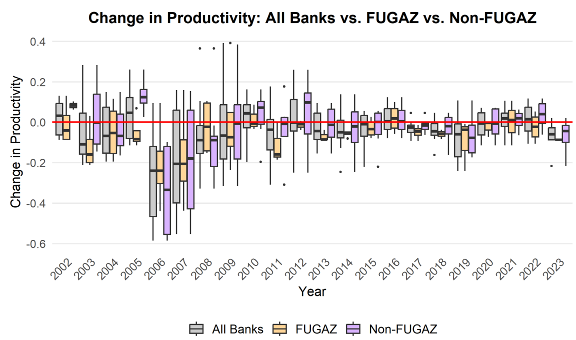
[10] The 2023 NLB average is based on the two nationally licensed banks with 2023 observations in the estimation sample. Both banks recorded productivity declines (−21.6% and −6.0%, respectively); the small cell size warrants caution in interpreting its magnitude.



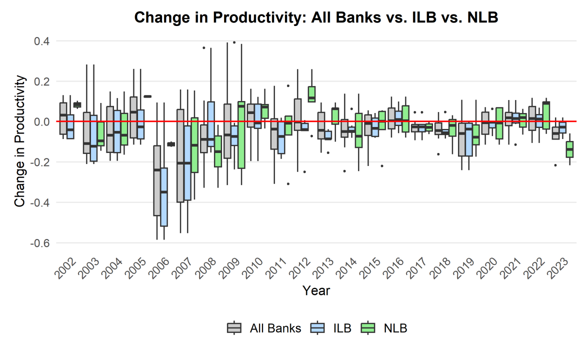

**Panel A:** Change in Productivity (All Banks, by Year)




**Panel B:** Change in Productivity Across Groups




**Panel C:** All Banks vs. FUGAZ vs. Non-FUGAZ




**Panel D:** All Banks vs. ILB vs. NLB

Figure 3: Change in productivity

Table 3: Mean and Median Productivity Change by Year and Subgroup

| | All | | ILB | | NLB | | FUGAZ | | Non-FUGAZ | |
|---|---|---|---|---|---|---|---|---|---|---|
| Year | Mean | Median | Mean | Median | Mean | Median | Mean | Median | Mean | Median |
| 2002 | 0.0210 | 0.0318 | -0.0097 | -0.0413 | 0.0826 | 0.0826 | -0.0097 | -0.0413 | 0.0826 | 0.0826 |
| 2003 | -0.0457 | -0.1088 | -0.0434 | -0.1225 | -0.0494 | -0.0950 | -0.1247 | -0.1593 | 0.0334 | -0.0018 |
| 2004 | -0.0385 | -0.0672 | -0.0463 | -0.0532 | -0.0280 | -0.0672 | -0.0463 | -0.0532 | -0.0280 | -0.0672 |
| 2005 | 0.0401 | 0.0467 | 0.0122 | -0.0264 | 0.1238 | 0.1238 | -0.0533 | -0.0834 | 0.1334 | 0.1238 |
| 2006 | -0.2725 | -0.2396 | -0.3259 | -0.3488 | -0.1125 | -0.1125 | -0.2063 | -0.2396 | -0.3388 | -0.3352 |
| 2007 | -0.1811 | -0.2054 | -0.2026 | -0.2054 | -0.1167 | -0.1167 | -0.1726 | -0.2054 | -0.1896 | -0.1791 |
| 2008 | -0.0666 | -0.0880 | 0.0103 | -0.0880 | -0.1564 | -0.1487 | -0.0239 | -0.0226 | -0.0856 | -0.0880 |
| 2009 | -0.0108 | -0.0667 | -0.0218 | -0.0730 | 0.0024 | 0.0756 | 0.0018 | -0.0730 | -0.0180 | -0.0069 |
| 2010 | 0.0272 | 0.0446 | -0.0056 | -0.0063 | 0.0601 | 0.0726 | 0.0157 | -0.0063 | 0.0322 | 0.0726 |
| 2011 | -0.0578 | -0.0370 | -0.0798 | -0.0742 | -0.0358 | -0.0079 | -0.1146 | -0.1580 | -0.0334 | -0.0079 |
| 2012 | 0.0237 | -0.0015 | -0.0674 | -0.0392 | 0.1149 | 0.1167 | -0.0141 | -0.0091 | 0.0399 | 0.0977 |
| 2013 | -0.0301 | -0.0434 | -0.0825 | -0.0853 | 0.0223 | 0.0642 | -0.0699 | -0.0853 | -0.0131 | -0.0130 |
| 2014 | -0.0415 | -0.0488 | -0.0315 | -0.0476 | -0.0536 | -0.0727 | -0.0575 | -0.0515 | -0.0324 | -0.0208 |
| 2015 | -0.0284 | -0.0088 | -0.0267 | -0.0334 | -0.0305 | 0.0019 | -0.0300 | -0.0334 | -0.0275 | 0.0019 |
| 2016 | 0.0157 | 0.0044 | 0.0165 | 0.0100 | 0.0148 | 0.0044 | 0.0198 | 0.0188 | 0.0133 | 0.0044 |
| 2017 | -0.0292 | -0.0268 | -0.0283 | -0.0268 | -0.0305 | -0.0285 | -0.0501 | -0.0463 | -0.0152 | -0.0153 |
| 2018 | -0.0420 | -0.0449 | -0.0416 | -0.0558 | -0.0426 | -0.0195 | -0.0583 | -0.0560 | -0.0312 | -0.0195 |
| 2019 | -0.0752 | -0.0586 | -0.0912 | -0.0366 | -0.0553 | -0.0769 | -0.0953 | -0.0366 | -0.0652 | -0.0769 |
| 2020 | -0.0113 | -0.0054 | -0.0087 | -0.0037 | -0.0139 | -0.0071 | -0.0252 | -0.0037 | -0.0054 | -0.0071 |
| 2021 | 0.0081 | 0.0187 | 0.0098 | 0.0153 | 0.0060 | 0.0203 | 0.0010 | 0.0119 | 0.0107 | 0.0195 |
| 2022 | 0.0161 | 0.0146 | -0.0020 | 0.0128 | 0.0413 | 0.0904 | -0.0249 | -0.0226 | 0.0365 | 0.0407 |
| 2023 | -0.0740 | -0.0595 | -0.0313 | -0.0273 | -0.1379 | -0.1379 | -0.0867 | -0.0867 | -0.0708 | -0.0434 |

*Notes:* Productivity change requires a valid one-period lag of recovered productivity, so this table is computed only from bank-year observations with finite $\Delta\hat{\omega}_{it}$. In 2002 and 2004, within this productivity-change subsample, the license and size partitions coincide: ILB observations are FUGAZ banks and NLB observations are Non-FUGAZ banks. Tables 4 and 5 use all finite RTS and TC observations for each year, because these level measures do not require a valid productivity-change lag; additional first-observed bank-year rows in 2002 and 2004 therefore break the same partition coincidence in those tables.

### 6.3 Returns to Scale

In a multi-output cost function, returns to scale (RTS) is defined as the reciprocal of the sum of cost elasticities with respect to output. Formally, RTS measures how the total cost responds to a proportional increase in all outputs. If a 1% increase in all outputs leads to a less than 1% increase in total cost (RTS>1) for a bank, it benefits from expanding production, indicating economies of scale. Given the duality of cost and IDF, we define RTS from the IDF as RTS $= \left(\sum_m \frac{\partial \ln X_1}{\partial \ln Y_{m,it}}\right)^{-1}$. Thus, if RTS exceeds 1, the bank operates under increasing returns to scale; its output bundle is sub-optimal, and there are economies of scale from expansion. The scale of operation is optimal when RTS $= 1$ (constant returns to scale). On the other hand, if RTS $< 1$ for a bank, it can lower its costs by reducing outputs. Note that RTS in our model is observation-specific.

### 6.3.1 Overall RTS Dynamics

Figure 4 and Table 4 present the estimated RTS for all banks over the period 2002–2023. The central result is that RTS exceeds unity throughout the sample, indicating that Nigerian banks operated under increasing returns to scale. The sample-wide mean RTS is 1.4625, and the median is 1.4652, so the typical bank remained below its cost-minimizing scale. RTS for every bank-year observation in the baseline estimation sample exceeds unity, with a sample minimum of 1.21.

Panel A of Figure 4 shows that RTS is persistent. Mean RTS was 1.3246 in 2002, 1.3287 in 2004, and 1.3420 in 2005. It rose during the post-consolidation and crisis years, reaching 1.5132 in 2009 and 1.5038 in 2010. After 2012, RTS remained in a relatively narrow band: 1.4859 in 2012, 1.4746 in 2013, 1.5196 in 2020, 1.5422 in 2022, and 1.5059 in 2023. The annual path therefore shows a stable and persistent gap from constant returns rather than a widening scale gap.

### 6.3.2 RTS and Recapitalization

Around the first recapitalization period, RTS remained moderately above unity. Mean RTS was 1.3287 in 2004, 1.3420 in 2005, and 1.3843 in 2006. This suggests that the first recapitalization did not immediately eliminate scale economies. Although consolidation may have helped banks expand their balance sheets, the sector still operated below its cost-minimizing scale.

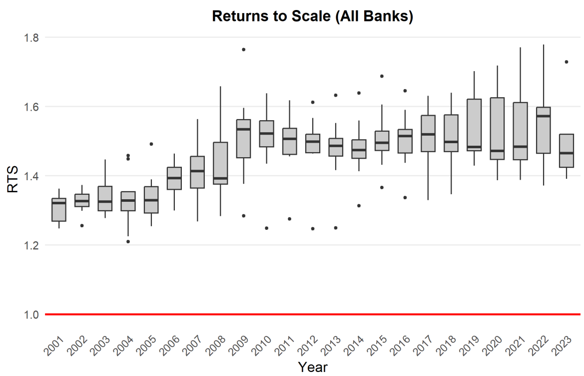
Following the post-2012 restructuring of the banking sector, RTS remained above unity but did not move sharply away from constant returns. The all-bank mean was 1.4859 in 2012 and 1.5059 in 2023. The post-2012 pattern therefore points to persistent unrealized economies of scale rather than a widening scale gap. A plausible interpretation is that banks continued to face constraints that prevented output expansion from fully exploiting their fixed and organizational capacity, including weak credit demand, asset-quality problems, regulatory restrictions, macroeconomic instability, exchange-rate pressures, and the rising cost of maintaining banking infrastructure.

### 6.3.3 Subgroup Heterogeneity

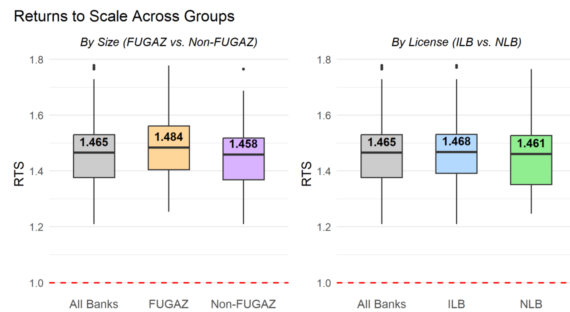
Panel B of Figure 4 summarizes the distribution of RTS across bank groups. The subgroup differences are modest. Along the size dimension, FUGAZ banks have a median RTS of 1.484, while Non-FUGAZ banks have a median RTS of 1.458. Along the license dimension, ILBs have a median RTS of 1.468, compared with 1.461 for NLBs. These values indicate that increasing returns to scale are a system-wide feature rather than a problem concentrated only among smaller or nationally licensed banks.

Panels C and D of Figure 4 show that the annual subgroup paths move closely together for much of the sample. In 2023, FUGAZ banks have a higher mean RTS of 1.7287, while Non-FUGAZ banks have a mean of 1.4502. Along the license dimension, ILBs record 1.5148 and NLBs 1.4926. These final-year values indicate that all groups remained under increasing returns to scale, but the differences are far smaller than under no trimming or winsorizing (see Section 6.6.8).

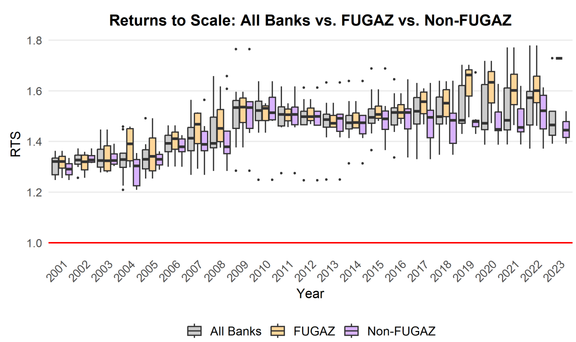
The persistent RTS values above unity carry two main implications. First, Nigerian banks continued to operate below their cost-minimizing scale throughout the sample period, meaning that output expansion could reduce average costs. Second, because subgroup gaps are modest, the scale inefficiency should be interpreted as a broad structural feature of the banking sector rather than as a problem confined to any single bank category. This helps explain the persistent productivity deterioration documented earlier: even when banks experience technical change or technological investment, failure to operate at an efficient scale can offset these gains and keep input costs elevated.



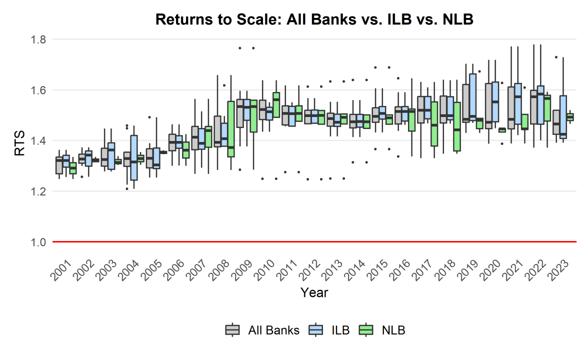

**Panel A:** RTS Over Time (All Banks)




**Panel B:** RTS Across Groups




**Panel C:** All Banks vs. FUGAZ vs. Non-FUGAZ




**Panel D:** All Banks vs. ILB vs. NLB

Figure 4: Returns to scale (RTS)

Table 4: Mean and Median Returns to Scale by Year and Subgroup

| | All | | ILB | | NLB | | FUGAZ | | Non-FUGAZ | |
|---|---|---|---|---|---|---|---|---|---|---|
| Year | Mean | Median | Mean | Median | Mean | Median | Mean | Median | Mean | Median |
| 2002 | 1.32456 | 1.32669 | 1.32572 | 1.34210 | 1.32264 | 1.31890 | 1.31395 | 1.32027 | 1.33517 | 1.32669 |
| 2003 | 1.33962 | 1.32541 | 1.35267 | 1.36233 | 1.31787 | 1.31342 | 1.34315 | 1.32442 | 1.33609 | 1.32541 |
| 2004 | 1.32868 | 1.32874 | 1.32866 | 1.31569 | 1.32872 | 1.32874 | 1.38443 | 1.39068 | 1.28408 | 1.30342 |
| 2005 | 1.34198 | 1.32952 | 1.33832 | 1.30405 | 1.35295 | 1.35295 | 1.35696 | 1.34148 | 1.32700 | 1.32952 |
| 2006 | 1.38426 | 1.39302 | 1.39061 | 1.39302 | 1.36205 | 1.36205 | 1.39643 | 1.41065 | 1.37453 | 1.37872 |
| 2007 | 1.41244 | 1.41338 | 1.40640 | 1.38889 | 1.42049 | 1.43957 | 1.43587 | 1.46837 | 1.40307 | 1.38889 |
| 2008 | 1.43979 | 1.39242 | 1.44186 | 1.40752 | 1.43736 | 1.37223 | 1.47407 | 1.45167 | 1.42455 | 1.37960 |
| 2009 | 1.51317 | 1.53432 | 1.51161 | 1.53132 | 1.51506 | 1.53432 | 1.51105 | 1.53612 | 1.51439 | 1.53432 |
| 2010 | 1.50380 | 1.52249 | 1.50211 | 1.51393 | 1.50549 | 1.56134 | 1.50539 | 1.53105 | 1.50312 | 1.51393 |
| 2011 | 1.49172 | 1.50671 | 1.50095 | 1.50608 | 1.48249 | 1.50733 | 1.50449 | 1.50608 | 1.48625 | 1.50733 |
| 2012 | 1.48588 | 1.49851 | 1.50364 | 1.49779 | 1.46812 | 1.49923 | 1.51089 | 1.49779 | 1.47516 | 1.49923 |
| 2013 | 1.47458 | 1.48680 | 1.47960 | 1.47213 | 1.46856 | 1.49113 | 1.48798 | 1.47213 | 1.46693 | 1.49113 |
| 2014 | 1.47788 | 1.47406 | 1.48020 | 1.47541 | 1.47510 | 1.47406 | 1.49077 | 1.47541 | 1.47051 | 1.47406 |
| 2015 | 1.50713 | 1.49519 | 1.51172 | 1.50643 | 1.50163 | 1.48966 | 1.52478 | 1.50643 | 1.49704 | 1.48966 |
| 2016 | 1.50401 | 1.51471 | 1.51428 | 1.51494 | 1.49169 | 1.51471 | 1.52229 | 1.51494 | 1.49357 | 1.51471 |
| 2017 | 1.50630 | 1.51981 | 1.53030 | 1.51981 | 1.47032 | 1.46047 | 1.54778 | 1.55721 | 1.47866 | 1.49533 |
| 2018 | 1.50438 | 1.49821 | 1.52897 | 1.49821 | 1.46750 | 1.44201 | 1.55179 | 1.55162 | 1.47277 | 1.48333 |
| 2019 | 1.53244 | 1.48345 | 1.56176 | 1.49561 | 1.50312 | 1.47983 | 1.62023 | 1.66285 | 1.49482 | 1.47849 |
| 2020 | 1.51959 | 1.47169 | 1.56261 | 1.55225 | 1.46797 | 1.44567 | 1.60784 | 1.63382 | 1.48650 | 1.44749 |
| 2021 | 1.52782 | 1.48386 | 1.56341 | 1.57296 | 1.47799 | 1.44629 | 1.60971 | 1.60177 | 1.48687 | 1.45543 |
| 2022 | 1.54217 | 1.57251 | 1.56012 | 1.58382 | 1.51703 | 1.56614 | 1.61030 | 1.60042 | 1.50810 | 1.52132 |
| 2023 | 1.50589 | 1.46546 | 1.51477 | 1.42448 | 1.49257 | 1.49257 | 1.72870 | 1.72870 | 1.45019 | 1.44497 |

## 6.4 Technical Change

Technical change (TC), derived from the input distance function, captures shifts in the production frontier over time. It is defined as $\text{TC}_{it} = \frac{\partial \ln D_{it}}{\partial t}$, evaluated at each observation. Since the IDF is dual to the cost function, a negative value of TC indicates technical progress, as the frontier shifts outward, enabling banks to produce the same output bundle at lower input cost. A positive value indicates technical regress, where costs rise for a given level of output. TC in our model is observation-specific, varying across banks and over time.

### 6.4.1 Overall TC Dynamics

Figure 5 and Table 5 present the estimated technical change for all banks over 2002 to 2023. The overall mean TC across the full sample is $-0.0351$ and the median is $-0.0354$, indicating a persistent outward shift of the production frontier.

Panel B of Figure 5 summarizes the distribution of TC across bank groups. The median is $-0.0354$ for all banks, $-0.0402$ for FUGAZ, $-0.0318$ for Non-FUGAZ, $-0.0386$ for ILB, and $-0.0304$ for NLB. Along both the size and license dimensions, larger and internationally licensed banks exhibit somewhat faster frontier advancement, but all groups show cost-reducing technical change on average.

#### 6.4.2 Pre-Recapitalization: 2002 to 2004

Technical progress was strongest in the early sample, with a mean TC of $-0.0563$ for all banks over 2002 to 2004. During this period, NLBs exhibited a faster pace of frontier advancement, averaging $-0.0608$, compared to $-0.0537$ for ILBs. Similarly, Non-FUGAZ banks averaged $-0.0606$ relative to $-0.0515$ for FUGAZ banks. These early patterns likely reflect the scope for efficiency gains among smaller and domestically focused banks operating in a highly fragmented pre-consolidation environment, where basic investments in operational systems could yield meaningful reductions in input costs.

#### 6.4.3 First Recapitalization

The mean TC for all banks during 2005 to 2011 was $-0.0467$, indicating continued technical progress following the consolidation exercise, although at a slower pace than in 2002–2004. ILBs averaged $-0.0495$ over 2005 to 2011, while NLBs averaged $-0.0437$. Along the size dimension, FUGAZ banks averaged $-0.0496$ compared to $-0.0451$ for Non-FUGAZ banks. The greater TC among FUGAZ and ILBs during this period is consistent with the greater capacity of larger banks to invest in electronic banking channels, risk-management systems, and back-office automation after consolidation.

#### 6.4.4 The Post-2012 Period

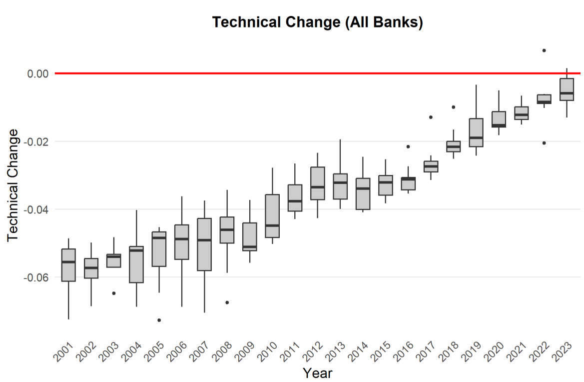
The post-2012 period is characterized by sustained but more moderate technical progress. For the full sample, mean TC over 2012 to 2023 is $-0.0222$, compared with $-0.0467$ during 2005–2011 and $-0.0563$ during 2002–2004. The pace of advancement therefore slows over time rather than deepens after 2012. Subgroup differences persist: over 2012–2023, FUGAZ banks average $-0.0242$ and ILBs average $-0.0247$, compared with $-0.0211$ for Non-FUGAZ and $-0.0191$ for NLBs. By 2023, the all-bank mean TC is $-0.0054$, indicating that technical progress remained cost-reducing but had become much smaller in magnitude.

#### 6.4.5 Discussion of TC Results

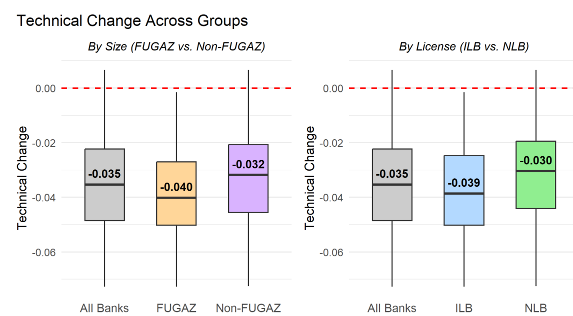
The persistent technical progress is consistent with a sustained period of technological and organizational change in the Nigerian banking sector. The strongest frontier shifts occur early in the sample, when the sector had large scope for basic modernization and operational catch-up. The 2004–2005 recapitalization and the post-2010 regulatory reforms, by strengthening bank balance sheets and consolidating the sector, likely enhanced the capacity of surviving institutions to invest in digital infrastructure, electronic banking channels, and back-office automation, but the marginal cost-reducing effect of these investments appears to have moderated over time.

The post-2012 slowdown in TC does not imply technological regress. TC remains negative for all groups on average, so the frontier continues to move outward. Rather, the estimates suggest that the largest gains from early automation and organizational modernization were realized before the later sample years, while subsequent digital investments delivered smaller marginal reductions in input requirements.

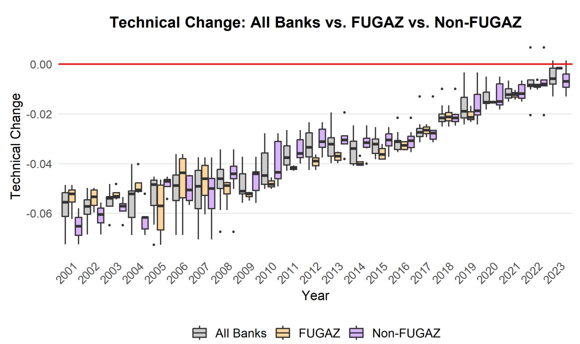
The contrast between sustained technical progress and negative productivity change underscores that technical progress alone does not guarantee productivity gains. Scale sub-optimality, macroeconomic volatility, asset-quality costs, and other operational frictions appear to have substantially offset the potential cost reductions implied by the shifting frontier throughout the sample period.



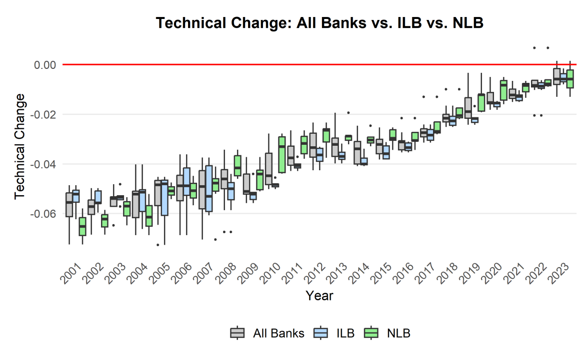

**Panel A:** TC Over Time (All Banks)




**Panel B:** TC Across Groups




**Panel C:** All Banks vs. FUGAZ vs. Non-FUGAZ




**Panel D:** All Banks vs. ILB vs. NLB

Figure 5: Technical change (TC)

Table 5: Mean and Median Technical Change by Year and Subgroup

| | All | | ILB | | NLB | | FUGAZ | | Non-FUGAZ | |
|---|---|---|---|---|---|---|---|---|---|---|
| Year | Mean | Median | Mean | Median | Mean | Median | Mean | Median | Mean | Median |
| 2002 | -0.0577 | -0.0572 | -0.0544 | -0.0557 | -0.0631 | -0.0623 | -0.0541 | -0.0534 | -0.0613 | -0.0604 |
| 2003 | -0.0552 | -0.0539 | -0.0532 | -0.0535 | -0.0584 | -0.0570 | -0.0522 | -0.0532 | -0.0581 | -0.0571 |
| 2004 | -0.0559 | -0.0522 | -0.0535 | -0.0514 | -0.0608 | -0.0614 | -0.0483 | -0.0506 | -0.0621 | -0.0616 |
| 2005 | -0.0534 | -0.0484 | -0.0542 | -0.0480 | -0.0509 | -0.0509 | -0.0583 | -0.0570 | -0.0485 | -0.0471 |
| 2006 | -0.0493 | -0.0488 | -0.0489 | -0.0488 | -0.0508 | -0.0508 | -0.0481 | -0.0437 | -0.0503 | -0.0506 |
| 2007 | -0.0505 | -0.0491 | -0.0504 | -0.0531 | -0.0507 | -0.0478 | -0.0475 | -0.0460 | -0.0518 | -0.0500 |
| 2008 | -0.0469 | -0.0461 | -0.0522 | -0.0500 | -0.0408 | -0.0416 | -0.0509 | -0.0489 | -0.0451 | -0.0441 |
| 2009 | -0.0487 | -0.0511 | -0.0518 | -0.0522 | -0.0451 | -0.0441 | -0.0527 | -0.0522 | -0.0465 | -0.0441 |
| 2010 | -0.0419 | -0.0448 | -0.0483 | -0.0484 | -0.0355 | -0.0331 | -0.0480 | -0.0484 | -0.0393 | -0.0435 |
| 2011 | -0.0364 | -0.0376 | -0.0405 | -0.0407 | -0.0322 | -0.0318 | -0.0417 | -0.0416 | -0.0340 | -0.0359 |
| 2012 | -0.0329 | -0.0334 | -0.0370 | -0.0364 | -0.0289 | -0.0264 | -0.0393 | -0.0389 | -0.0302 | -0.0312 |
| 2013 | -0.0326 | -0.0322 | -0.0365 | -0.0370 | -0.0279 | -0.0288 | -0.0373 | -0.0370 | -0.0299 | -0.0304 |
| 2014 | -0.0346 | -0.0339 | -0.0387 | -0.0401 | -0.0297 | -0.0304 | -0.0396 | -0.0403 | -0.0318 | -0.0315 |
| 2015 | -0.0325 | -0.0321 | -0.0352 | -0.0359 | -0.0292 | -0.0297 | -0.0357 | -0.0363 | -0.0307 | -0.0304 |
| 2016 | -0.0313 | -0.0312 | -0.0332 | -0.0334 | -0.0289 | -0.0306 | -0.0328 | -0.0327 | -0.0304 | -0.0308 |
| 2017 | -0.0264 | -0.0274 | -0.0281 | -0.0284 | -0.0237 | -0.0269 | -0.0266 | -0.0266 | -0.0262 | -0.0277 |
| 2018 | -0.0205 | -0.0216 | -0.0220 | -0.0227 | -0.0182 | -0.0205 | -0.0210 | -0.0212 | -0.0202 | -0.0216 |
| 2019 | -0.0173 | -0.0189 | -0.0215 | -0.0217 | -0.0131 | -0.0121 | -0.0205 | -0.0214 | -0.0159 | -0.0187 |
| 2020 | -0.0133 | -0.0153 | -0.0161 | -0.0155 | -0.0100 | -0.0083 | -0.0151 | -0.0153 | -0.0126 | -0.0149 |
| 2021 | -0.0115 | -0.0122 | -0.0131 | -0.0128 | -0.0094 | -0.0084 | -0.0123 | -0.0123 | -0.0111 | -0.0118 |
| 2022 | -0.0078 | -0.0084 | -0.0099 | -0.0087 | -0.0049 | -0.0077 | -0.0085 | -0.0088 | -0.0075 | -0.0080 |
| 2023 | -0.0054 | -0.0058 | -0.0051 | -0.0058 | -0.0058 | -0.0058 | -0.0015 | -0.0015 | -0.0063 | -0.0069 |

### 6.5 Reserve for Loan Losses Ratio

We incorporate the ratio of reserve for loan losses to gross loans (RFLL) as an environmental variable into the input distance function framework. RFLL captures differences in asset quality across banks and over time, since higher loan-loss reserves reflect greater expected credit losses and weaker loan portfolio quality. Including RFLL allows us to adjust productivity estimates for differences in banks' credit-risk exposure, so that measured productivity is not confounded with deterioration in loan quality.

The economic interpretation of RFLL follows from the duality between the input distance function and the cost function. Because the IDF is dual to the cost function, the estimated marginal effect of RFLL in the IDF translates into a cost effect: $\partial \ln C/\partial \text{RFLL} = \partial \ln D/\partial \text{RFLL}$, holding outputs and input prices fixed.[11] A higher RFLL is expected to raise costs because banks with weaker asset quality must devote more resources to credit monitoring, loan recovery, loan restructuring, legal procedures, provisioning, and write-off management.

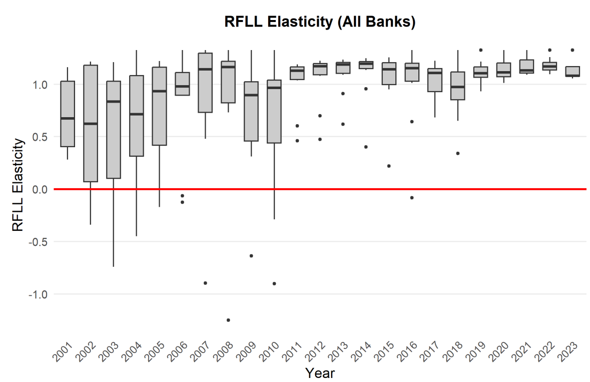
Panel A of Figure 6 displays the estimated RFLL cost effect over time for all banks. The estimates indicate a positive central tendency, with a mean of 0.9395 and a median of 1.1108 across all bank-year

[11] RFLL is a ratio, so it enters in levels rather than logs. A one percentage-point increase corresponds to a 0.01 increase in the RFLL variable.

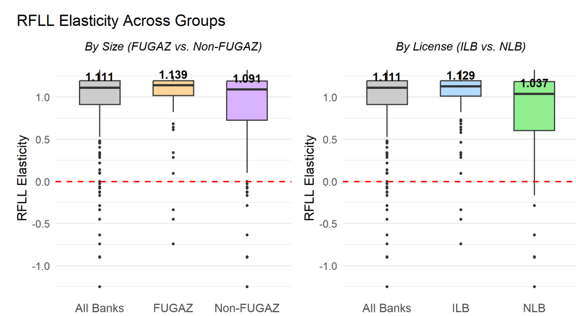
observations. Since RFLL is measured as a ratio, these magnitudes imply that a one percentage-point increase in RFLL is associated with approximately a 0.94% increase in cost at the mean and a 1.11% increase at the median, holding outputs and input prices constant.

Panel B of Figure 6 shows that the cost effect of RFLL is broadly similar across bank groups, with moderate differences across the two decompositions. The median cost effect is 1.139 for FUGAZ banks, 1.091 for Non-FUGAZ banks, 1.129 for ILBs, and 1.037 for NLBs. The ordering indicates that larger and internationally licensed banks bear a somewhat heavier cost burden from loan-loss provisioning, plausibly because of more complex corporate, sovereign, and cross-border exposures, but the positive cost effect is common to all groups.

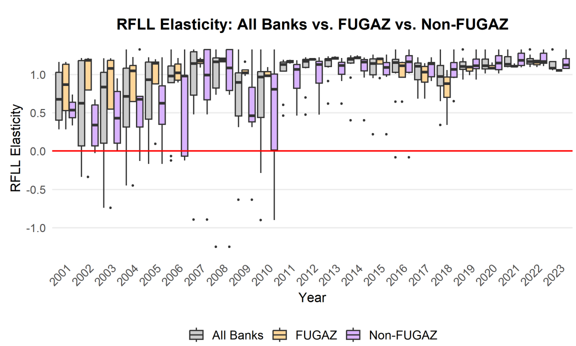
Panels C and D of Figure 6 show the evolution of the RFLL cost effect over time by subgroup. The mean effect rises from roughly 0.55–0.75 in the early 2000s to values above one in much of the post-2011 period. This indicates that asset-quality deterioration became more costly as the banking system became larger, more complex, and more tightly regulated after consolidation. Overall, the results suggest that weaker asset quality consistently raises banks' costs, with the effect being somewhat stronger among FUGAZ and ILB institutions.



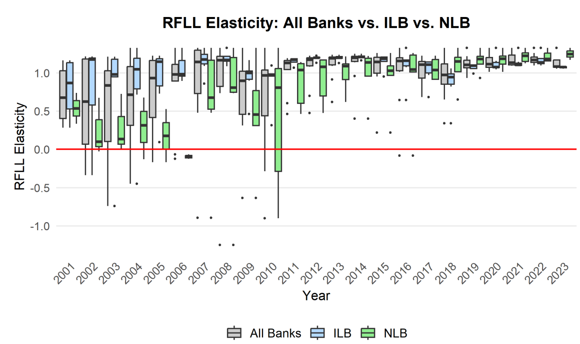

**Panel A:** RFLL Cost Elasticity Over Time (All Banks)




**Panel B:** RFLL Cost Elasticity Across Groups




**Panel C:** All Banks vs. FUGAZ vs. Non-FUGAZ




**Panel D:** All Banks vs. ILB vs. NLB

Figure 6: RFLL Cost elasticity

## 6.6 Robustness Checks

We subject our baseline estimates to a series of robustness checks to assess the stability and validity of our findings. The results are summarized in Tables 6, 7, 8, and 9, and in Figure 7.

### 6.6.1 Baseline Instrument Validity

We estimate the baseline model using a nonlinear two-step GMM procedure. The instruments are constructed to be orthogonal to the innovation component in the productivity process, $\xi_{it}$. The baseline specification yields a Hansen $J$-statistic of 3.7322 with 17 degrees of freedom and a $p$-value of 0.9997, so we do not reject the null hypothesis that the overidentifying restrictions hold. However, the unusually large $p$-value is not stronger evidence of instrument validity. With 43 moment conditions and a small bank panel, it may reflect limited test power or dependence among the moment conditions. We therefore interpret the test cautiously and place greater weight on the stability of the findings across alternative instrument sets, functional forms, and restricted samples.

### 6.6.2 Instrument Relevance

Although our control function approach does not require a formal first stage as in Olley & Pakes (1996), Levinsohn & Petrin (2003), and Ackerberg et al. (2015), we assess instrument relevance by regressing the dependent variable $\ln X_{1,it}$ on the full set of variables entering the moment conditions — the lagged dependent variable, the contemporaneous and one-period-lagged translog terms, the lagged RFLL ratio and its square, and lagged total assets — together with bank fixed effects. Accordingly, the model yields an $F(42, 173) = 185.35$ ($p < 0.001$), and the regression has an adjusted $R^2$ of 0.984. The variables underlying the moment conditions therefore have strong explanatory power for $\ln X_{1,it}$.

### 6.6.3 Alternative Estimator: NLS

We re-estimate the model by nonlinear least squares (NLS) via the `nl2sol` trust-region algorithm (the `port` routine in R's `nls`) as an alternative to GMM. The estimated persistence parameter $\hat{\rho}_1$ is 0.7257, and the mean change in productivity is $\overline{\%\Delta\hat{\omega}} = -3.9\%$, consistent in sign and direction with the baseline. This confirms that the finding of declining bank productivity is not an artifact of the GMM estimator.

### 6.6.4 Exclusion of the Crisis Period and Crisis–Size Interaction

To assess sensitivity to the global financial crisis, we re-estimate the baseline model after dropping observations from 2008 to 2010. Because the crisis indicator is identically zero in the restricted sample, its coefficient is not estimated. The specification retains the baseline set of 43 moment conditions, including the two lagged reserve-ratio moments, and estimates 25 parameters; hence, the Hansen test has $43 - 25 = 18$ degrees of freedom. The resulting estimation sample contains 185 bank-year observations.

The estimated persistence parameter is $\hat{\rho}_1 = 0.7372$, and the $J$-test yields a statistic of 7.0402 with a $p$-value of 0.9898. Mean and median productivity changes are $-2.6\%$ and $-2.3\%$, respectively, so the baseline productivity pattern does not hinge on the crisis years.

As a further check on the crisis specification, we augment the 2008–2010 crisis dummy in (2) with an interaction between the crisis dummy and lagged total assets, $GFC_{it} \times \ln G_{it-1}$, to allow the crisis effect on productivity to differ by bank size. In the interaction specification, the crisis dummy itself is negative but imprecisely estimated ($\hat{\delta}_{\text{GFC}} = -0.4438$, bank-clustered $p = 0.2968$). The interaction coefficient is positive, $\hat{\delta}_{GFC\times G} = 0.0327$, with a bank-clustered standard error of 0.0312 and a $p$-value of 0.2948. The negative stand-alone dummy is not evidence of instability relative to the baseline estimate of 0.0373 in Section 6.2.4: in the interaction model the dummy measures the crisis effect for a hypothetical bank with $\ln G_{it-1} = 0$, far outside the data. Evaluated at the sample mean of lagged log total assets (13.37), the implied crisis-period shift is $-0.4438 + 0.0327 \times 13.37 \approx -0.01$, which is economically negligible and statistically indistinguishable from the small, insignificant baseline coefficient. Thus, neither the crisis dummy nor the crisis-size interaction is statistically significant, i.e., we find no evidence that the crisis-period productivity shift varied systematically with bank size. The Hansen test again does not reject the overidentifying restrictions ($J = 4.3646$, $p = 0.9996$, 18 degrees of freedom), and the substantive results are unchanged, with mean and median productivity changes of $-3.4\%$ and $-3.1\%$, mean and median RTS of 1.4992 and 1.4776, and mean and median technical change of $-0.0350$ and $-0.0352$. Although imprecisely estimated, the positive interaction point estimate implies a slightly smaller crisis-period productivity penalty for larger banks, directionally consistent with the subgroup pattern in Section 6.2.4, where FUGAZ and internationally licensed banks were relatively more resilient during 2008–2010. Panel B of Table 8 reports the full interaction diagnostics. Allowing the crisis effect to vary with size therefore preserves the baseline conclusion of declining productivity, cost-reducing technical change, and persistent increasing returns to scale.

### 6.6.5 Alternative Functional Form: Cobb–Douglas

We replace the translog input distance function with a Cobb–Douglas specification to assess sensitivity to functional form. The Cobb–Douglas estimates yield a persistence parameter of $\hat{\rho}_1 = 0.6882$ and a mean productivity change of $\overline{\%\Delta\hat{\omega}} = -5.1\%$. Of 14 candidate moments, the lagged time trend is collinear with the intercept and contemporaneous time trend and is removed, leaving 13 retained moments. With 10 estimated parameters, the $J$-test therefore has $13 - 10 = 3$ degrees of freedom. The statistic of 4.9542 ($p = 0.1752$) fails to reject instrument validity at the 5% level. The direction and magnitude of productivity change are consistent with the more flexible translog specification.

### 6.6.6 Choice of Alternative Instrument Sets

We further assess robustness to instrument choice by re-estimating the baseline model under alternative instrument sets. Excluding selected input-ratio instruments and the lag of total assets yields a mean productivity change of $-3.4\%$, a Hansen $J$-statistic of 4.4265 ($p = 0.9923$, 14 degrees of freedom), and a persistence parameter of 0.7266. The baseline estimate of declining productivity is therefore not driven by the inclusion of any particular instrument and is robust to plausible variation in the instrument set.

### 6.6.7 Sample Restrictions: NLB-only and Excluding the Largest Banks

A sector-wide productivity-decline claim should not be driven by a single bank group or by the largest institutions in the sample. We therefore report two additional sample-restriction checks. Both restricted-sample models retain the baseline productivity process, including the lagged-total-assets shifter $\ln G_{it-1}$. First, we exclude internationally licensed banks and re-estimate the model using only the seven nationally licensed banks. The NLB-only restricted estimation sample contains 90 bank-year observations. The NLB-only GMM estimate has $\hat{\rho}_1 = 0.1999$ and $\hat{\delta}_G = -0.3560$ (bank-clustered standard error 0.1135, $p = 0.0017$). The overidentification test supports the moment restrictions ($J = 17.7257$, $p = 0.4063$, 17 degrees of freedom). Mean and median productivity changes are $-2.9\%$ and $-2.6\%$, respectively. Thus, the average decline persists when a separate technology is estimated for nationally licensed banks.

Second, we exclude the five FUGAZ banks and re-estimate the model on the remaining 12 banks, directly addressing whether the full-sample decline is driven by the largest tier-one institutions. The Excluding-FUGAZ restricted estimation sample contains 141 bank-year observations. The estimate has $\hat{\rho}_1 = 0.2661$ and $\hat{\delta}_G = -0.1819$ (bank-clustered standard error 0.0661, $p = 0.0059$). The $J$-test again fails to reject the overidentifying restrictions ($J = 14.6414$, $p = 0.6213$, 17 degrees of freedom). Mean productivity declines by 2.3%, while the median change is approximately zero (0.03%). The negative mean after removing FUGAZ shows that the average productivity decline is not generated solely by the five largest banks, although the near-zero median points to heterogeneity among the remaining institutions. Together, the two restrictions support a sector-wide decline in mean productivity while indicating that its magnitude and distribution vary across bank groups.

Table 6: Mean Productivity Change: Sample-Restriction Robustness Checks

| **Specification** | **Obs.** | $\hat{\rho}_1$ | Mean $\%\Delta\hat{\omega}$ | Median $\%\Delta\hat{\omega}$ | $J$-statistic ($df$) | $p$-value |
|---|---|---|---|---|---|---|
| Baseline (17 banks) | 230 | 0.7375 | −3.5 | −3.1 | 3.732 (17) | 0.9997 |
| NLB-only (7 banks) | 90 | 0.1999 | −2.9 | −2.6 | 17.726 (17) | 0.4063 |
| Excluding-FUGAZ (12 banks) | 141 | 0.2661 | −2.3 | 0.0 | 14.641 (17) | 0.6213 |
| GFC × lagged size interaction | 230 | 0.7232 | −3.4 | −3.1 | 4.365 (18) | 0.9996 |

*Notes:* Obs. denotes estimation observations after trimming, lag construction, and complete-case restrictions. Mean and median productivity changes are reported in percent. NLB-only excludes all internationally licensed banks; in this sample, it also excludes all FUGAZ banks. Excluding-FUGAZ removes the five largest tier-one banks and re-estimates a separate technology on the remaining 12 banks. The GFC × lagged size row is a full-sample crisis-interaction robustness check. The Hansen degrees of freedom equal the number of retained moment conditions minus estimated parameters. Baseline, NLB-only, and Excluding-FUGAZ each retain 43 moments and estimate 26 parameters, including the lagged-total-assets productivity shifter $\delta_G$, giving 17 overidentifying restrictions. The crisis-size interaction uses 45 moments and 27 parameters.

Table 7: Mean and Median Productivity Change (in percent) across several Robustness Specifications

| **Specification** | Mean | Median |
|---|---|---|
| Baseline Translog GMM | −3.5 | −3.1 |
| Crisis Period Exclusion (GMM) | −2.6 | −2.3 |
| GFC × Lagged Size Interaction (GMM) | −3.4 | −3.1 |
| Alternative Estimator (NLS) | −3.9 | −3.3 |
| Cobb–Douglas Functional Form (GMM) | −5.1 | −4.2 |
| Alternative Instrument Sets (GMM) | −3.4 | −2.3 |
| NLB-only Re-estimation (GMM) | −2.9 | −2.6 |
| Excluding-FUGAZ Banks (GMM) | −2.3 | 0.0 |

*Notes:* $\%\Delta\hat{\omega}_{it} = (\hat{\omega}_{it} - \hat{\omega}_{it-1}) \times 100$ denotes the percentage change in productivity. Mean and median are reported in percent. The GFC × lagged size row augments the full-sample productivity process with the interaction $GFC_{it} \times \ln G_{it-1}$. The NLB-only and Excluding-FUGAZ rows estimate separate technologies on restricted samples and are therefore sample-composition robustness checks rather than common-technology subgroup statistics.

Table 8: Diagnostic Statistics Across Robustness Checks

*Panel A: Common robustness diagnostics*

| | **Crisis Excl.** | **NLS** | **Cobb–Douglas** | **Alt. Instruments** | **NLB-only GMM** | **Excl. FUGAZ** |
|---|---|---|---|---|---|---|
| $\hat{\rho}_1$ | 0.7372 | 0.7257 | 0.6882 | 0.7266 | 0.1999 | 0.2661 |
| $J$-statistic | 7.0402 | — | 4.9542 | 4.4265 | 17.7257 | 14.6414 |
| $J$-test $p$-value | 0.9898 | — | 0.1752 | 0.9923 | 0.4063 | 0.6213 |
| Degrees of freedom | 18 | — | 3 | 14 | 17 | 17 |

*Panel B: GFC × lagged-size interaction diagnostics*

| **Statistic** | **Estimate** |
|---|---|
| Estimation observations | 230 |
| $\hat{\rho}_1$ | 0.7232 |
| $\hat{\delta}_{GFC}$ | $-0.4438$ |
| $se(\hat{\delta}_{GFC})$ | 0.4253 |
| $p(\hat{\delta}_{GFC})$ | 0.2968 |
| $\hat{\delta}_{GFC\times G}$ | 0.0327 |
| $se(\hat{\delta}_{GFC\times G})$ | 0.0312 |
| $p(\hat{\delta}_{GFC\times G})$ | 0.2948 |
| $J$-statistic | 4.3646 |
| $J$-test $p$-value | 0.9996 |
| Degrees of freedom | 18 |
| Mean / median productivity change (%) | $-3.4$ / $-3.1$ |
| Mean / median RTS | 1.4992 / 1.4776 |
| Mean / median technical change | $-0.0350$ / $-0.0352$ |

*Notes:* $\hat{\rho}_1$ is the estimated AR(1) persistence coefficient. $\hat{\delta}_{GFC\times G}$ is the coefficient on $GFC_{it} \times \ln G_{it-1}$. The $J$-statistic is Hansen's overidentification test. The $J$-test is not applicable for NLS. In the crisis-exclusion specification, the GFC coefficient is omitted because the indicator has no variation; all 43 baseline moments are retained and 25 parameters are estimated, giving 18 overidentifying restrictions. The Cobb–Douglas specification retains 13 moments after removing the collinear lagged time trend and estimates 10 parameters, giving 3 overidentifying restrictions. The degrees of freedom for the other checks likewise equal their retained moment conditions minus their estimated parameters.

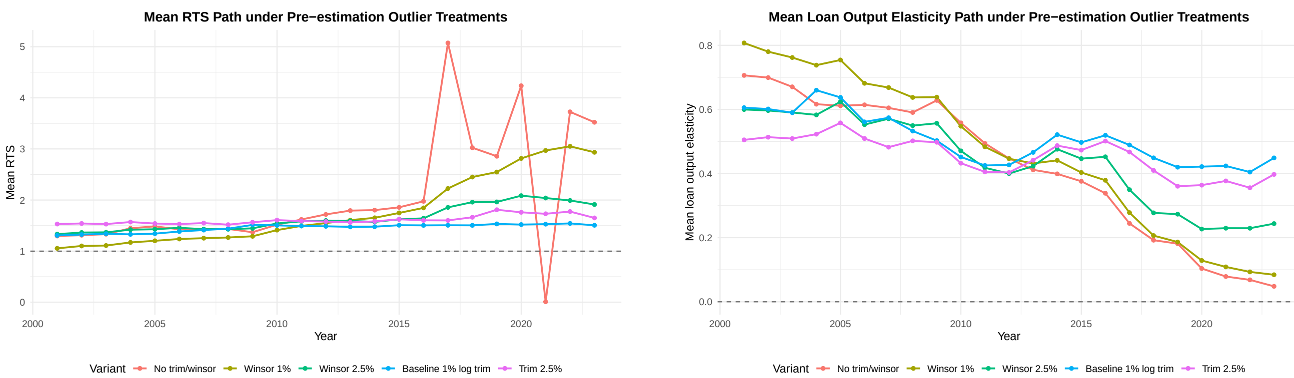
#### 6.6.8 Outlier Treatment

Finally, we examine whether the RTS results are driven by extreme observations in the logged variables. The baseline model uses data with outliers trimmed at the top and bottom 1% level prior to estimation. Table 9 reports this baseline alongside the "No trimming or winsorizing," winsorized, and stronger-trimmed sensitivity variants, while Figure 7 shows the corresponding yearly RTS and loan-elasticity paths. The "No trimming or winsorizing" specification produces a mean RTS of 2.064, compared with a mean RTS of 1.462 and a median RTS of 1.465 under the baseline 1% logged-variable trim, with no baseline observations above an RTS of two. As an external yardstick, Das & Kumbhakar (2012) report an all-bank RTS of 1.0802 for Indian banks. Our trimmed Nigerian estimates remain above unity, but they are much closer to this banking-literature benchmark than the estimate with no trimming or winsorizing. Thus, the conclusion that Nigerian banks operate under increasing returns is robust, but the extreme untrimmed RTS values are substantially attenuated once outliers are handled before estimation.

Table 9: Outlier Robustness of Scale Estimates: Data Trimming Before GMM Estimation

| Variant | N | $\rho_1$ | Mean RTS | Median RTS | Mean $\epsilon_1$ | Share RTS > 2 | $J$ ($p$) |
|---|---|---|---|---|---|---|---|
| No trimming or winsorizing | 254 | 0.725 | 2.064 | 1.654 | 0.408 | 0.343 | 8.49 (0.9550) |
| winsorize 1% | 254 | 0.742 | 1.820 | 1.550 | 0.449 | 0.303 | 7.37 (0.9783) |
| winsorize 2.5% | 254 | 0.717 | 1.652 | 1.602 | 0.433 | 0.114 | 5.12 (0.9974) |
| Baseline 1% trim (logged variables) | 230 | 0.738 | 1.462 | 1.465 | 0.501 | 0.000 | 3.73 (0.9997) |
| trim 2.5% | 212 | 0.741 | 1.611 | 1.592 | 0.453 | 0.042 | 5.52 (0.9959) |

*Notes:* The baseline row is the main estimation reported in the paper: the logged variables are trimmed at the top and bottom 1% before estimation. The "No trimming or winsorizing" row applies neither treatment; the 1% and 2.5% winsorized rows and the 2.5% trimmed row are sensitivity variants.




**Panel A:** RTS Path

**Panel B:** Loan Elasticity Path

Figure 7: Outlier Robustness Paths

# 7 Discussion of Results

The empirical results reveal a “scale-technology paradox” in the Nigerian banking sector. The industry shows persistent cost-reducing technical change, yet aggregate productivity declines — with steeper losses among the largest and internationally licensed banks, but declines in mean productivity also present in the restricted samples — because banks operate below their cost-minimizing scale and continue to face operating frictions associated with lending, asset quality, and macroeconomic volatility. We focus on these dynamics using an IDF estimated via nonlinear GMM on an unbalanced panel of 17 commercial banks over 2000–2023.

The economic intuition is straightforward. Technology can reduce the amount of inputs needed to produce a given bundle of loans and investments, but those savings are not automatic. They become realized productivity gains only when banks can expand sound lending, use existing organizational capacity more fully, and keep credit-risk costs under control. In Nigeria, weak borrower information, collateral-enforcement problems, macroeconomic instability, and loan-loss pressures limit this translation from technical progress to measured productivity.

## 7.1 The Scale-Technology Paradox and Productivity Erosion

Despite consistent technical progress, evidenced by a mean technical change (TC) of $-0.0351$ and a median of $-0.0354$, overall productivity declined by an average of 3.5% annually, with a median decline of 3.1%. The subgroup estimates show steeper losses among the largest and internationally licensed banks, but the restricted-sample mean declines of 2.9% for NLB-only banks and 2.3% after excluding FUGAZ show that the pattern is not confined to them. This divergence indicates that technological gains have been offset by scale sub-optimality and other operating frictions.

Nigerian banks operated under increasing returns to scale throughout the sample, with a sample-wide mean RTS of 1.46 and a median of 1.47. The annual path is stable: mean RTS is 1.325 in 2002 and 1.506 in 2023. The sector therefore does not converge to constant returns. The coexistence of negative productivity change and technical progress suggests that while the frontier of Nigerian banking is advancing through modernization, banks remain below efficient scale, generating higher-than-necessary input costs that erode the benefits of progress.

## 7.2 Cost Elasticities and the Dominance of Lending

Our analysis of output cost elasticities confirms that loan production remains the primary expense driver.

- Loan vs. Investment Intensity: The median loan elasticity (0.490) is more than twice the median investment elasticity (0.202), implying that a 1% increase in loan output requires about a 0.49% increase in inputs, compared to about 0.20% for investments.

- Asset Quality Costs: The reserve for loan losses (RFLL) enters the IDF with a positive cost effect. Since RFLL is measured as a ratio, the mean coefficient of 0.940 implies that a one percentage-point increase in the RFLL ratio is associated with roughly a 0.94% increase in cost. The median effect is about 1.11%.

- Input Substitution: The implied input ratio elasticities for labor and capital are approximately $-2.99$ and $-4.11$, indicating that both are substitutes for deposits and borrowings, with stronger substitutability for capital.

### 7.3 Heterogeneity: FUGAZ vs. Non-FUGAZ Banks

Partitioning the sample reveals that productivity losses are larger among the sector's leaders, even though increasing returns to scale are present across all groups. FUGAZ and ILB banks exhibit somewhat faster technical progress, but they also experience more severe productivity declines: mean productivity changes are $-5.6\%$ for FUGAZ banks and $-5.0\%$ for ILBs, compared with $-2.3\%$ for Non-FUGAZ banks and $-1.5\%$ for NLBs. RTS differences across groups are modest, with median values clustered between 1.46 and 1.48. The main heterogeneity therefore lies more in productivity and technical-change dynamics than in radically different scale positions.

### 7.4 Robustness and Statistical Validity

The finding of declining productivity is robust in the baseline common-technology model. Hansen's $J$-test does not reject the overidentifying restrictions. The unusually large $p$-value is not stronger evidence of instrument validity and may partly reflect limited test power or the number and dependence of the moment conditions relative to the small bank panel. We therefore place greater weight on the consistency of the findings across alternative instrument sets, functional forms, and restricted samples. The auxiliary instrument-relevance regression has an $F$-statistic of 185.35, indicating strong explanatory power; this diagnostic alone does not establish instrument validity. The finding of declining productivity remains robust across several alternative specifications:

- Alternative Estimators: Nonlinear Least Squares (NLS) estimation yields a mean productivity change of $-3.9\%$.

- Crisis Sensitivity: Excluding the 2008–2010 global financial crisis period results in mean and median productivity changes of $-2.6\%$ and $-2.3\%$, respectively.

- Crisis-Size Interaction: Allowing the GFC effect to vary with lagged total assets gives a mean productivity change of $-3.4\%$, and the interaction coefficient is not statistically significant ($p = 0.295$).

- Functional Form: Replacing the translog with a Cobb–Douglas specification yields a productivity change of $-5.1\%$.

- Alternative Instruments: A restricted instrument set yields a productivity change of $-3.4\%$ and a Hansen $J$-test $p$-value of 0.992.

- Sample Restrictions: Estimating a separate NLB-only technology yields mean and median productivity declines of 2.9% and 2.6%. Excluding the five FUGAZ banks yields a mean decline of 2.3% and an approximately zero median change. Because both restricted models retain the baseline lagged-size productivity shifter, these results show that the decline in mean productivity is not driven solely by internationally licensed or FUGAZ banks, while the near-zero Non-FUGAZ median indicates meaningful heterogeneity among smaller institutions.

Outlier checks further show that the scale conclusion is robust to trimming or winsorizing the variables before estimation.

### 7.5 Policy Implications

Our findings carry several policy implications for the Nigerian banking sector and its regulators.

First, the persistent increasing returns to scale indicate that banks would benefit from expanding productive output relative to their fixed and organizational capacity. Because subgroup RTS differences are modest, this is a system-wide issue rather than a problem limited to smaller NLB or Non-FUGAZ institutions. Regulatory policies that facilitate sound credit expansion, improve borrower information, support collateral enforcement, and deepen underserved credit markets could help banks move closer to their cost-minimizing scale. The updated estimates suggest that the 2004–2005 capital-driven consolidation and the post-2012 restructuring of the sector did not deliver convergence to unitary RTS: mean RTS was 1.486 in 2012 and 1.506 in 2023. Further M&A-style consolidation alone is therefore unlikely to resolve scale inefficiency unless it is accompanied by better deployment of balance-sheet capacity into productive intermediation.

Second, technical progress is a persistent and robust feature of the sector, but its magnitude moderates over time. Mean TC is $-0.056$ during 2002–2004, $-0.047$ during 2005–2011, and $-0.022$ during 2012–2023. This pattern suggests that early modernization and consolidation generated larger technology shifts, while later digital investments continued to reduce costs but with smaller marginal effects. Policy should therefore support not only technology adoption, but also the organizational and market conditions that allow technological gains to translate into productivity gains.

Third, the finding that productivity has declined despite consistent technical progress highlights the importance of addressing operational and structural factors that prevent banks from translating frontier improvements into actual cost reductions. These include scale sub-optimality, weak credit-market infrastructure, macroeconomic instability, high funding costs, and asset-quality deterioration.

Fourth, loans remain substantially more cost-intensive than investments across the banking system. The implication is that system-wide improvements to credit-market infrastructure would yield disproportionate cost savings in the activity that accounts for the largest share of intermediation. Strengthening credit bureaus, improving collateral registries, modernizing the legal framework for secured transactions, and streamlining judicial processes for loan recovery would lower the cost of financial intermediation across the system.

Finally, the mean reserve-for-loan-losses cost effect of 0.94 implies that a one percentage-point increase in the RFLL ratio is associated with approximately a 0.94% increase in cost at the mean, while the median effect is approximately 1.11%. This finding underscores that prudential supervision of asset quality has direct cost-side benefits in addition to its more familiar stability-side rationale. Strengthening NPL resolution mechanisms, refining loan classification and provisioning standards, supporting bank-level capacity for credit monitoring, and ensuring that frameworks such as AMCON remain effective at clearing legacy bad loans would directly reduce intermediation costs, particularly during periods of macroeconomic stress.

# 8 Conclusion

This study provides the first application of the control function approach to estimate banking productivity in Nigeria using an input distance function framework. Our analysis reveals that Nigerian banks experienced a decline in productivity from 2000 to 2023, despite persistent cost-reducing technical change. The primary explanation for this apparent paradox is that banks have not reached their optimal scale of operations: returns to scale remain above unity throughout the sample, and the resulting scale inefficiency offsets part of the cost savings from technological progress. The baseline decline is stable across alternative estimators, functional forms, crisis-period exclusion, crisis-size interaction, and alternative instruments. Mean productivity also declines in the NLB-only and Excluding-FUGAZ re-estimations, showing that the pattern is not confined to larger and internationally active banks, although those groups record steeper losses in the baseline subgroup comparisons. The results underscore the need for a comprehensive policy approach that combines continued support for technological adoption with measures to facilitate sound output expansion, improve credit-market infrastructure, strengthen asset-quality supervision, and address the structural barriers that prevent banks from fully realizing their economies of scale. A more productive banking sector will strengthen financial intermediation and, through it, support the growth of the broader Nigerian economy.

## Declaration of Generative AI and AI-Assisted Technologies in the Manuscript Preparation Process

During the preparation of this work, the authors used OpenAI Codex to assist with code verification and language editing. All substantive decisions regarding code writing, estimation methods, interpretation, and presentation of results were made by the authors. After using this tool, the authors reviewed and edited the code, output, and manuscript as needed, and take full responsibility for the content of the publication.

## Data Availability

The bank-level dataset used in this study was obtained through Bloomberg Terminal and is subject to licensing restrictions; therefore, it cannot be publicly redistributed. However, the underlying variables can be reconstructed from the publicly available annual reports and financial statements of the sampled banks. To facilitate reproducibility, the replication package provides the complete data-processing and estimation code, a synthetic dataset that reproduces the structure of the restricted data, and detailed replication instructions. The replication package is available through Mendeley Data at `https://doi.org/10.17632/w5f32znz38.1`.

## References


Ackerberg, D. A., Caves, K., & Frazer, G. (2015). Identification properties of recent production function estimators. *Econometrica*, 83(6), 2411–2451.

Ahamed, M. M., Ho, S. J., Mallick, S. K., & Matousek, R. (2021). Inclusive banking, financial regulation and bank performance: Cross-country evidence. *Journal of Banking & Finance*, 124, 106055.

Ahamed, M. M. & Mallick, S. K. (2017). Does regulatory forbearance matter for bank stability? evidence from creditors' perspective. *Journal of Financial Stability*, 28, 163–180.

Al-Azzam, M., Bernstein, D. H., & Parmeter, C. F. (2026). Cost efficiency and gender diversity in microfinance: A stochastic frontier approach. *Economic Modelling*, 156, 107473.

Apeti, A. E., Bambe, B.-W.-W., & Lompo, A. A. B. (2024). Determinants of public sector efficiency: A panel database from a stochastic frontier analysis. *Oxford Economic Papers*, 76(3), 741–758.

Asmild, M. & Matthews, K. (2012). Multi-directional efficiency analysis of efficiency patterns in Chinese banks 1997–2008. *European Journal of Operational Research*, 219(2), 434–441.

Barros, C. P., Managi, S., & Matousek, R. (2012). The technical efficiency of the Japanese banks: Non-radial directional performance measurement with undesirable output. *Omega*, 40(1), 1–8.

Benston, G. J. (1965). Branch banking and economies of scale. *The Journal of Finance*, 20(2), 312–331.

Berger, A. N. (2007). International comparisons of banking efficiency. *Financial Markets, Institutions & Instruments*, 16(3), 119–144.

Berger, A. N., Hancock, D., & Humphrey, D. B. (1993). Bank efficiency derived from the profit function. *Journal of Banking & Finance*, 17(2–3), 317–347.

Berger, A. N., Hanweck, G. A., & Humphrey, D. B. (1987). Competitive viability in banking: Scale, scope, and product mix economies. *Journal of Monetary Economics*, 20(3), 501–520.

Berger, A. N. & Humphrey, D. B. (1997). Efficiency of financial institutions: International survey and directions for future research. *European Journal of Operational Research*, 98(2), 175–212.

Central Bank of Nigeria (2000). *Banking Supervision Annual Report 2000*. Technical report, Bank Examination and Banking Supervision Departments, Central Bank of Nigeria, Abuja, Nigeria. ISSN 1595-0387. Universal banking discussed in Section 1.03, p. 5; Guidelines for the Practice of Universal Banking in Nigeria (Circular BSD/DO/CIR/VOL.1/10/2000, 22 December 2000) reproduced as Appendix 1, pp. 86–94.

Central Bank of Nigeria (2010). Regulation on the scope of banking activities & ancillary matters, no. 3, 2010. Circular No. BSD/DIR/RBG/2010/08. Effective November 15, 2010.

Central Bank of Nigeria (2024). *Financial Stability Report: December 2024*. Technical report, Central Bank of Nigeria, Abuja, Nigeria.

Creel, J., Hubert, P., & Labondance, F. (2023). Credit, banking fragility, and economic performance. *Oxford Economic Papers*, 75(2), 553–573.

Das, A. & Kumbhakar, S. C. (2012). Productivity and efficiency dynamics in Indian banking: An input distance function approach incorporating quality of inputs and outputs. *Journal of Applied Econometrics*, 27(2), 205–234.

Delis, M. D. & Tsionas, E. G. (2009). The joint estimation of bank-level market power and efficiency. *Journal of Banking & Finance*, 33(10), 1842–1850.

Färe, R. & Primont, D. (1995). *Multi-Output Production and Duality: Theory and Applications*. Boston, MA: Kluwer Academic Publishers.

Feng, J., Jiao, W., & Wang, Y. (2025). Bank fintech and lending distance: The role of information asymmetry and risk management channels. *Economic Modelling*, 151, 107154.

Fethi, M. D. & Pasiouras, F. (2010). Assessing bank efficiency and performance with operational research and artificial intelligence techniques: A survey. *European Journal of Operational Research*, 204(2), 189–198.

Fukuyama, H., Matousek, R., & Tzeremes, N. G. (2025). Detecting skimping in bank production process with non-performing loans: A distance minimization approach. *Economic Modelling*, 152, 107305.

International Monetary Fund (2025). World Economic Outlook Database, April 2025: Gross domestic product, deflator – Nigeria. World Economic Outlook database. WEO subject code NGDP_D. Units: index, 2010 = 100. Accessed July 6, 2026.

International Monetary Fund (2026a). Geographical outreach: Number of automated teller machines (ATMs), country wide for Nigeria [ngafcacnum]. Retrieved from FRED, Federal Reserve Bank of St. Louis. Financial Access Survey. Accessed February 3, 2026.

International Monetary Fund (2026b). Geographical outreach: Number of branches, excluding headquarters, for commercial banks for Nigeria [ngafcbodcnum]. Retrieved from FRED, Federal Reserve Bank of St. Louis. Financial Access Survey. Accessed February 3, 2026.

International Monetary Fund (2026c). Geographical outreach: Number of commercial banks for Nigeria [ngafciodcnum]. Retrieved from FRED, Federal Reserve Bank of St. Louis. Financial Access Survey. Accessed February 3, 2026.

International Monetary Fund (2026d). Use of financial services: Key indicators, outstanding deposits with commercial banks for Nigeria (percent of GDP) [ngafclodcggdppt]. Retrieved from FRED, Federal Reserve Bank of St. Louis. Financial Access Survey. Accessed February 3, 2026.

Kumbhakar, S. C. (2013). Specification and estimation of multiple output technologies: A primal approach. *European Journal of Operational Research*, 231(2), 465–473.

Levinsohn, J. & Petrin, A. (2003). Estimating production functions using inputs to control for unobservables. *The Review of Economic Studies*, 70(2), 317–341.

Mallick, S., Rughoo, A., Tzeremes, N. G., & Xu, W. (2020). Technological change and catching-up in the Indian banking sector: A time-dependent nonparametric frontier approach. *Computational Economics*, 56, 217–237.

Nigeria Deposit Insurance Corporation (2024). *Financial Condition and Performance of Deposit Money Banks in the Third and Fourth Quarter, 2024*. NDIC Quarterly Journal Vol. 39, Nos. 3–4, Nigeria Deposit Insurance Corporation, Abuja, Nigeria.

Olley, G. S. & Pakes, A. (1996). The dynamics of productivity in the telecommunications equipment industry. *Econometrica*, 64(6), 1263–1297.

Osuagwu, E. S., Isola, W. A., & Nwaogwugwu, I. C. (2018). Measuring technical efficiency and productivity change in the Nigerian banking sector: A comparison of non-parametric and parametric techniques. *African Development Review*, 30(4), 490–501.

Sanusi, S. L. (2010). The nigerian banking industry: What went wrong and the way forward. Convocation lecture delivered at Bayero University, Kano, February 26, 2010. Central Bank of Nigeria. Accessed January 28, 2025.

Tsionas, M. G. & Mallick, S. K. (2019). A Bayesian semiparametric approach to stochastic frontiers and productivity. *European Journal of Operational Research*, 274, 391–402.

Wang, Y., Kumbhakar, S. C., & Jin, M. (2025). The yin and yang of banking: Modeling desirable and undesirable outputs. *European Journal of Operational Research*, 322(3), 1025–1044.

Wu, X. & Wu, X. (2025). How does bank digitalization benefit firm borrowing? Evidence from China. *Economic Modelling*, 153, 107313.

Zhao, T. & Murinde, V. (2011). Bank deregulation and performance in nigeria. *African Development Review*, 23(1), 30–43.