

Network Productivity and Loan-Income Adjustment after Monetary Policy Normalization: Evidence from Japanese Regional Banks

Hirohisa Shimura

shimura@soka.ac.jp

Soka University

Research Article

Keywords: Japanese regional banks, monetary policy normalization, network DEA, lending intermediation, loan interest income, Malmquist productivity index

Posted Date: September 25th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10859122/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Japan's exit from negative interest rates offers a setting to examine how bank intermediation adjusts to monetary policy normalization. We analyze 96 Japanese regional banking entities from FY2019 to FY2025 using a multi-period additive relational network DEA model, in which deposits become loans in a lending stage and loans generate loan interest income in an earnings stage. Productivity change is evaluated mainly through the Global Malmquist index and its efficiency-change/technical-change decomposition; a stage-level diagnostic, reducing algebraically to loan-to-deposit and loan-yield growth rates, supplements this evidence.

The two stage-level ratios were broadly aligned before normalization. The gap stayed small in FY2023–FY2024 but widened substantially in FY2024–FY2025, mainly because the loan-yield ratio rose more than the loan-to-deposit ratio, in 91 of 96 entities. Because it relates loan interest income to loans, this increase reflects stronger relative loan-income performance rather than an independent productivity result, and may capture repricing and portfolio composition as well as operational change. Regional differences were limited initially but became significant in the second year; Regional Banks I and II showed almost identical gaps.

Global productivity, measured by the Malmquist index, increased in both post-change comparisons, with a larger rise in the loan-yield ratio than the loan-to-deposit ratio, especially in the second comparison, varying across regions. The results show where measured change concentrated within the intermediation network but do not separate repricing from changes in production technology; comparisons are descriptive rather than causal.

JEL Classification: C61; D24; E52; G21

1. Introduction

Japan's move away from negative interest rates changed the environment in which regional banks had operated for many years. In March 2024, the Bank of Japan (BOJ) ended the negative interest rate policy and yield curve control and returned to a framework centered on the short-term interest rate (Bank of Japan, 2024a). The change matters for regional banks because deposits, loans, and the interest income earned on those loans remain central to their business. Higher lending rates can support interest income, but higher rates can also affect funding costs, credit demand, and competition among banks (Borio et al., 2017; Claessens et al., 2018).

These effects do not necessarily reach all parts of bank intermediation at the same time. Loan volumes depend on credit demand, borrower relationships, local economic conditions, and banks' lending decisions. Interest income also depends on the rates charged on new loans and the repricing of existing loans. The BOJ's Financial System Report noted that long-term lending rates had begun to rise after the change in the interest-rate environment, while some short-term lending rates were expected to adjust

with a lag (Bank of Japan, 2024b). These differences in timing mean that lending activity and the income generated from lending may not adjust at the same pace.

This possibility is important for Japanese regional banks. They provide much of the financing used by local firms and small and medium-sized enterprises, while operating in markets affected by demographic decline, weak credit demand in some areas, and prolonged pressure on interest margins (Hoshi & Kashyap, 2004; Borio et al., 2017; Claessens et al., 2018). Studies of Japanese banks have documented considerable differences in efficiency and productivity across banks and over time (Fukuyama, 1995; Fukuyama & Matousek, 2017). Much of this evidence, however, tells us how the bank performs as a whole. It tells us less about whether lending and the income generated from lending change in the same way.

Network data envelopment analysis (DEA) allows the two activities to be examined within the same production process. In this study, deposits are used to produce loans, and loans are then used to generate loan interest income. Loans are therefore both the output of the first stage and the input of the second. The first stage captures lending intermediation, while the second captures earnings intermediation. This structure preserves the connection between the two activities without assuming that their performance must move together (Tone & Tsutsui, 2009; Fukuyama & Matousek, 2017).

Kourtzidis et al. (2021) apply a multi-period network DEA model to Japanese Regional Banks I and II from 2002 to 2017. Their approach makes it possible to compare network performance across years using a common multi-period structure. It also provides evidence of differences in productivity and efficiency among Japanese regional banks. Their sample ends well before the BOJ's 2024 policy change, however. Whether lending activity and relative loan-income performance continue to move together when interest rates begin to normalize remains an open question.

We examine this question using 96 Japanese regional banking entities observed from FY2019 to FY2025. In this study, FY t denotes the Japanese fiscal year from April 1 of calendar year t through March 31 of $t + 1$. The Bank of Japan's March 19, 2024 policy decision therefore occurred 12 days before FY2023 ended: the FY2023 observation is overwhelmingly pre-change, whereas FY2024 (April 1, 2024–March 31, 2025) is the first complete post-change fiscal year and FY2025 is the second. Four annual transitions (FY2019–FY2020 through FY2022–FY2023) form the pre-normalization benchmark; FY2023–FY2024 compares an almost entirely pre-change fiscal year with the first complete post-change fiscal year, and FY2024–FY2025 compares the first and second complete post-change fiscal years. Because all sampled banks were exposed to the national policy environment and no contemporaneous untreated domestic group is available, these are descriptive, event-aligned before-after comparisons rather than causal treatment estimates.

Before normalization, changes in lending and earnings performance were generally close. The gap was still relatively small in FY2023–FY2024 but widened substantially in FY2024–FY2025. The lending-stage change index remained above one in the second year, while the earnings-stage index increased much more strongly. Because the earnings stage uses loans and loan interest income, this pattern may

capture differences in loan repricing and portfolio composition in addition to differences in operational performance. The widening occurred across most banks, although its magnitude differed across regions. Regional Banks I and Regional Banks II, in contrast, showed little difference in their post-normalization stage gaps.

The distinction between lending and earnings is important for interpreting these changes. Overall productivity alone does not show where within the intermediation process the change occurred. It also does not show that the difference between the two stages became much larger only in the second year after normalization.

This study offers three related contributions. The first is a measurement contribution: within a multi-period additive relational network DEA model built on a single common multiplier set, the stage-specific common-weight change indices reduce algebraically to the year-on-year change in two ordinary financial ratios—the loan-to-deposit ratio and loan interest income per unit of loans—so that the Global Malmquist productivity index and its efficiency-change/technical-change decomposition, rather than the stage indices themselves, constitute the primary DEA-derived productivity evidence in this setting. The second is an empirical contribution: applying this framework to a constant-perimeter panel of 96 Japanese regional banking entities spanning the BOJ's 2024 policy normalization, we document that overall productivity rose and that the loan-to-deposit and loan-yield ratios diverged more sharply in the second post-change year than in the first, a pattern that a whole-bank productivity measure alone would not reveal. The third is a methodological contribution that is also a limitation: because the ratio reduction follows directly from the single-input, single-output specification of each stage, the same reduction will arise in any two-stage network DEA application built on this structure, so the approach is best understood as a transparent way of organizing ratio-based diagnostics within a common-weight productivity framework, generalizable to other single-input/single-output network settings, rather than as a new DEA efficiency construct in its own right.

2. Literature Review and Research Positioning

2.1 Bank Intermediation and Network DEA

DEA has been widely used to measure bank efficiency without imposing a specific functional form on the production technology (Charnes et al., 1978; Banker et al., 1984), while productivity change over time can be measured using the Malmquist productivity index (Färe et al., 1994). Conventional DEA models, however, generally treat the bank as a single production unit.

This treatment can obscure the internal structure of bank intermediation. Deposits are used to fund loans, but loans also generate interest income. Loans therefore have a dual role: they are an output of lending activity and an input into income generation. Two-stage and network DEA models make this relationship explicit by linking production stages through intermediate products (Chen & Zhu, 2004; Chen et al., 2009, 2010; Tone & Tsutsui, 2009).

This distinction matters when lending and the income generated from lending do not change at the same pace. Loan volumes depend on credit demand and banks' lending decisions, whereas loan interest income also reflects lending rates and the repricing of existing loans. Fukuyama and Matousek (2017) show how network DEA can represent such internal banking activities within a connected production process.

Changes in the interest-rate environment make this distinction particularly relevant. Loan balances may change little even as the income generated from those loans rises. An overall productivity measure combines these movements and does not show where within the intermediation process the change occurs. A two-stage network structure allows lending and earnings performance to be examined separately while retaining the link between them.

2.2 Japanese Regional Banks and Monetary Policy Normalization

Japanese regional banks depend heavily on deposits and lending, particularly lending to small and medium-sized firms in their local markets. Demographic decline, weak credit demand in some regions, and competition among banks have placed pressure on this business model. Prolonged low interest rates added to these pressures by narrowing the margins available from traditional lending activities.

Low interest rates can compress net interest margins and reduce bank profitability, particularly when opportunities to expand lending or other sources of revenue are limited (Borio et al., 2017; Claessens et al., 2018). Japanese banks have also shown substantial differences in efficiency and productivity across banks and over time (Fukuyama, 1995). More recent work continues to show that bank performance can differ across components of the production process. Fukuyama et al. (2022), for example, examine several sources of inefficiency in Japanese bank production, while Shi et al. (2025) model banking performance through a multi-stage network structure.

The move away from very low interest rates changes this setting. Regional banks entered the new environment with loan portfolios accumulated during the preceding low-rate period. Rates on new loans can respond relatively quickly, while existing loans are repriced according to their contractual terms and reset schedules. Changes in loan volumes and loan interest income may therefore occur at different speeds.

In March 2024, the Bank of Japan ended its negative interest rate policy and changed the monetary-policy framework. March 19, 2024 is used here as a transparent timing marker for organizing the annual comparisons, not as an isolated exogenous treatment that identifies a causal effect. Because the decision occurred only 12 days before the March 31 close of FY2023, the FY2023-FY2024 transition contrasts an almost entirely pre-change fiscal year with the first complete post-change fiscal year.

2.3 Multi-Period Network Productivity in Japanese Regional Banking

Kourtzidis et al. (2021) develop a multi-period additive relational network DEA model and apply it to Japanese Regional Banks I and II over 2002–2017. Their model evaluates observations from different periods within a common production structure and provides overall and period-specific efficiency measures. It also uses common multipliers to measure productivity change across periods.

Their application documents differences in efficiency and productivity across banks, periods, and the two conventional categories of Japanese regional banks. More importantly for the present study, the multi-period structure makes it possible to follow changes within a connected banking production process without estimating each year against an unrelated technology.

The question examined here is different from that addressed by Kourtzidis et al. (2021). Their sample ends several years before the BOJ began monetary policy normalization. The 2024 policy change provides a setting in which the two connected stages of bank intermediation may adjust differently and in which that difference can be followed over successive years.

We use the multi-period network framework to examine this adjustment over time. Rather than asking only whether regional-bank productivity changed after normalization, we compare change in lending-stage and loan-income-stage performance and examine whether the difference between them changed as normalization progressed.

2.4 Research Gap and Positioning

Existing studies provide a basis for examining lending and earnings as connected parts of bank intermediation, but they say less about how the relationship between the two changes when the interest-rate environment shifts. The question is not only whether the two stages respond differently, but also when such a difference becomes visible.

This timing matters after a long period of very low interest rates. A change in rates need not appear immediately or equally in both stages. Looking only at the first year after a policy change may therefore give a different picture from one that follows the adjustment for a longer period.

There is also a question of where these differences appear. Earlier studies of Japanese regional banks have reported differences between Regional Banks I and Regional Banks II, but banks also operate in local markets with different credit demand and competitive conditions. The conventional bank-type classification and geographic location may therefore capture different aspects of post-normalization adjustment.

This study compares lending-stage and loan-income-stage performance before and during the first two years of monetary policy normalization. We examine whether the two stages changed at different rates, whether the difference became more pronounced as normalization progressed, and whether the pattern varied more across geographic regions or between Regional Banks I and Regional Banks II.

3. Methodology

3.1 Two-Stage Intermediation Structure

The production process is divided into two stages. Deposits are used to produce loans in Stage 1, and loans are used to generate loan interest income in Stage 2. This treatment follows the network DEA literature, where an intermediate product links one stage of production to the next (Chen & Zhu, 2004; Chen et al., 2009, 2010; Fukuyama & Matousek, 2017).

For bank j in fiscal year k :

$$x_j^{(k)} = \text{deposits}$$

1

$$z_j^{(k)} = \text{loans}$$

2

$$y_j^{(k)} = \text{loaninterestincome}$$

3

Stage 1 is defined as:

$$x_j^{(k)} \rightarrow z_j^{(k)}$$

4

Stage 2 is defined as:

$$z_j^{(k)} \rightarrow y_j^{(k)}$$

5

Loans are the output of Stage 1 and the input of Stage 2. The two stages are evaluated within the same network rather than as separate production processes.

The model focuses on the deposit-loan-loan-interest-income channel. Other activities of regional banks, such as securities investment and fee income, are not included. The purpose is to separate changes in lending from changes in the income generated by the loan portfolio. This minimal three-variable structure is a deliberate design choice rather than a data-availability shortcut. A single-input, single-output specification of each stage keeps the two quantities that the model aggregates transparent and directly interpretable: the loan-to-deposit ratio, as an intermediation-quantity measure, and loan interest income per unit of loans, as an intermediation-revenue-intensity measure—precisely the ratios that

Section 3.5 shows the stage-change indices reduce to. Adding further inputs or outputs, such as operating costs, risk-weighted assets, or fee income, would yield genuine multi-dimensional efficiency scores that do not collapse to simple ratios, but it would also mix loan-income growth with revenue from unrelated activities and would require cost, risk, and fee-income data that are not uniformly disclosed at the required granularity across all 96 sampled entities over FY2019–FY2025. The present specification therefore trades a richer efficiency construct for a transparent, fully replicable measure of how revenue generated per unit of the loan portfolio evolved relative to the underlying deposit base, which is the object motivated by the research question in Section 1.

3.2 Multi-Period Additive Relational Network DEA

To compare network performance across periods, we use the multi-period additive relational network DEA model of Kourtzidis et al. (2021). The model extends the additive relational two-stage framework to multiple periods by applying a common set of multipliers to observations from all periods.

For banking entity j in fiscal year k , let $x_j^{(k)}$, $z_j^{(k)}$, and $y_j^{(k)}$ denote deposits, loans, and loan interest income, respectively. For the banking entity under evaluation, o , v_o , w_o , and u_o denote the corresponding multipliers. Under constant returns to scale (CRS), the multi-period model is specified as follows: The objective and common-technology constraints are given in Equations (6)-(10).

$$\max_{(u_o, v_o, w_o)} \left[w_o \sum_{(k=1)}^K z_o^{(k)} + u_o \sum_{(k=1)}^K y_o^{(k)} \right]$$

6

subject to

$$v_o \sum_{(k=1)}^K x_o^{(k)} + w_o \sum_{(k=1)}^K z_o^{(k)} = 1$$

7

$$w_o z_j^{(k)} - v_o x_j^{(k)} \leq 0, j = 1, \dots, n; k = 1, \dots, K$$

8

$$u_o y_j^{(k)} - w_o z_j^{(k)} \leq 0, j = 1, \dots, n; k = 1, \dots, K$$

9

$$u_o, v_o, w_o \geq \epsilon$$

10

Here, n denotes the number of banking entities and the number of fiscal years. The same set of multipliers is applied to all observations across the study period. This allows each banking entity to be evaluated

against a common multi-period technology and provides a consistent basis for comparing its performance over time. The multi-period analysis is conducted under CRS. A VRS extension would require additional free variables and could reduce discriminatory power in the multi-period setting (Kourtzidis et al., 2021). In the numerical implementation, the non-Archimedean infinitesimal was represented by the strictly positive lower bound $\varepsilon = 10^{-8}$ on each multiplier. This value prevents zero multipliers while remaining negligible relative to the normalized data and objective values.

3.3 Period- and Stage-Specific Efficiency

Once the common optimal multipliers (v_o^* , w_o^* , u_o^*) have been obtained, lending and earnings efficiency can be evaluated separately for each fiscal year. These measures are defined in Equations (11)-(13).

The lending-stage efficiency of entity o in year k is

$$E_{L,o}^{(k)} = \frac{(w_o^* z_o^{(k)})}{(v_o^* x_o^{(k)})}$$

11

and the earnings-stage efficiency is

$$E_{E,o}^{(k)} = \frac{(u_o^* y_o^{(k)})}{(w_o^* z_o^{(k)})}$$

12

Both measures range from zero to one, with higher values indicating greater efficiency relative to the corresponding best-practice frontier.

The overall network efficiency for year k is

$$E_{N,o}^{(k)} = \frac{(w_o^* z_o^{(k)} + u_o^* y_o^{(k)})}{(v_o^* x_o^{(k)} + w_o^* z_o^{(k)})}$$

13

Using the same optimal multipliers across periods and stages preserves the network structure of the model while allowing changes in lending and earnings performance to be examined separately over time.

Because stage-specific efficiency scores may depend on the selected multiplier solution when multiple optimal solutions exist, we also checked whether the optimal multipliers were numerically unique. For each of the 96 banking entities, let E_o^* denote the optimal overall objective value and δ the numerical tolerance. Six additional linear programs were then solved to determine the minimum and maximum feasible values of each multiplier: The multiplier-range program is reported in Equations (14)-(19).

$$\frac{\min}{\max_{(v_o, w_o, u_o)} v_o}, \frac{\min}{\max_{(v_o, w_o, u_o)} w_o}, \frac{\min}{\max_{(v_o, w_o, u_o)} u_o} \text{ to (14)}$$

$$E_o^* - \delta \leq w_o \sum_{(k=1)}^K z_o^{(k)} + u_o \sum_{(k=1)}^K y_o^{(k)} \leq E_o^* + \delta$$

15

$$v_o \sum_{(k=1)}^K x_o^{(k)} + w_o \sum_{(k=1)}^K z_o^{(k)} = 1$$

16

$$w_o z_j^{(k)} - v_o x_j^{(k)} \leq 0, j = 1, \dots, n; k = 1, \dots, K$$

17

$$u_o y_j^{(k)} - w_o z_j^{(k)} \leq 0, j = 1, \dots, n; k = 1, \dots, K$$

18

$$u_o, v_o, w_o \geq \epsilon$$

19

The numerical tolerance δ was set to 10^{-8} . The resulting ranges were negligible for all 96 entities. The largest ranges observed in the sample were approximately 7.50×10^{-8} for v , 6.94×10^{-8} for w , and 8.78×10^{-8} for u . The common optimal multipliers were therefore used directly to calculate period- and stage-specific efficiencies.

3.4 Global Malmquist Productivity Change

Intertemporal productivity change is evaluated using the common-weight global Malmquist productivity index proposed in the multi-period framework of Kourtzidis et al. (2021). Let $E_o^G(t)$ denote the overall efficiency of entity o in year t relative to the common multi-period technology. Productivity change between years t and $t + 1$ is defined as This definition is reported in Eq. (20).

$$MPI_o^G(t, t + 1) = E_o^{G(t+1)} / E_o^{G(t)}$$

20

The interpretation is conventional: $MPI_o^{G(t,t+1)} > 1$ indicates productivity improvement, $MPI_o^{G(t,t+1)} = 1$ indicates no productivity change, and $MPI_o^{G(t,t+1)} < 1$ indicates productivity deterioration. Because all periods are evaluated relative to the same multi-period technology using a common multiplier system, the global index provides a consistent basis for comparing productivity across adjacent periods.

To distinguish changes in relative efficiency from shifts in the frontier, the global MPI is decomposed following the global Malmquist framework of Pastor and Lovell (2005). Let $E_o^C(t)$ denote the overall efficiency of entity o relative to the contemporaneous constant-returns-to-scale (CRS) frontier in year t . For each year, this efficiency is obtained by solving the single-period version of the same additive relational network DEA model, using only the n banking entities observed in that year. For focal entity o , the contemporaneous program is: The contemporaneous program is given in Equations (21)-(25).

$$E_o^C(t) = \max w_o^t z_o^t + u_o^t y_o^t$$

21

$$\text{subject to } v_o^t x_o^t + w_o^t z_o^t = 1$$

22

$$w_o^t z_j^t - v_o^t x_j^t \leq 0, \quad j = 1, \dots, n$$

23

$$u_o^t y_j^t - w_o^t z_j^t \leq 0, \quad j = 1, \dots, n$$

24

$$u_o^t, v_o^t, w_o^t \geq \epsilon$$

25

The optimal objective value from this year-specific program is $E_o^C(t)$. The program is solved independently for every entity and every year; observations from other years do not enter the contemporaneous reference set. The global efficiencies $E_o^G(t)$, by contrast, are evaluated against the common multi-period technology defined in Section 3.3.

Efficiency change between t and $t + 1$ is defined as This measure is shown in Eq. (26).

$$EC_o^{(t,t+1)} = E_o^C(t+1) / E_o^C(t)$$

26

Technical change, equivalently the change in the best-practice gap between the contemporaneous and global technologies, is defined as This measure is shown in Eq. (27).

$$TC_o^{(t,t+1)} = \left[E_o^G(t+1) / E_o^C(t+1) \right] / \left[E_o^G(t) / E_o^C(t) \right]$$

27

This definition ensures the exact multiplicative decomposition shown in Eq. (28).

$$MPI_o^{G(t,t+1)} = EC_o^{(t,t+1)} \times TC_o^{(t,t+1)}$$

28

Efficiency change captures movement toward or away from the year-specific CRS frontier. Technical change captures whether that contemporaneous frontier moves closer to or farther from the common multi-period frontier along the observed entity's ray. Values above one indicate improvement or technical progress, whereas values below one indicate deterioration or technical regress.

3.5 Stage-Specific Productivity Change and the Stage Gap

The main analysis focuses on how productivity change differs between the two connected stages.

Using the common-weight stage efficiencies defined in Section 3.3, lending-stage change is calculated as in Eq. (29).

$$C_{L,o}^{(t,t+1)} = E_{L,o}^{G(t+1)} / E_{L,o}^{G(t)}$$

29

while earnings-stage change is in Eq. (30).

$$C_{E,o}^{(t,t+1)} = E_{E,o}^{G(t+1)} / E_{E,o}^{G(t)}$$

30

For both measures, values above one indicate improvement, values equal to one indicate no change, and values below one indicate deterioration.

These ratios are referred to as common-weight stage change indices rather than separate Malmquist productivity indices. They are calculated from the stage efficiencies generated by the common multi-period multiplier solution and are not independently estimated Malmquist indices.

Because a single common multiplier triple is applied to each banking entity in every fiscal year, the two stage-change indices simplify algebraically rather than remaining independent efficiency measures. In Eq. (29), the optimal Stage 1 multipliers enter the numerator and denominator of the lending-stage

efficiency ratio identically in both years and cancel, so the lending-stage change index is identical to the year-on-year growth rate of the entity's loan-to-deposit ratio. The same cancellation applies in Eq. (30): the optimal Stage 2 multipliers cancel, so the earnings-stage change index is identical to the year-on-year growth rate of loan interest income per unit of loans, that is, loan yield. This reduction follows directly from the single-input, single-output specification of each stage under a common per-entity multiplier system; it holds for any two-stage network model built on this structure and is not specific to the present application. It does not indicate an error in the estimation, but it does mean that the two stage-change indices carry no information beyond the corresponding financial ratios: the multi-period DEA procedure supplies a single, theoretically motivated common-weight system under which the two ratios are compared and combined with the overall network efficiency and Global MPI measures, but the stage-level indices themselves do not constitute independent DEA productivity evidence.

To measure the relative movement of the two stages, we define the stage gap as in Eq. (31).

$$Gap_o^{(t,t+1)} = C_{E,o}^{(t,t+1)} - C_{L,o}^{(t,t+1)}$$

31

A positive gap indicates that earnings intermediation improved more, or deteriorated less, than lending intermediation during the transition. A negative gap indicates that lending intermediation improved more, or deteriorated less, than earnings intermediation.

Given the reduction established above, the stage gap is best understood as a ratio decomposition consistent with the network structure rather than as an alternative DEA efficiency or productivity index. It packages two economically meaningful, network-consistent financial ratios—loan-to-deposit growth and loan-yield growth—under the common evaluation frame shared with the overall network efficiency and Global MPI measures, and it is used as an auxiliary, diagnostic statistic to locate where within the connected production process aggregate productivity change concentrated. The primary DEA-derived evidence in this study is instead the Global Malmquist productivity index and its efficiency-change/technical-change decomposition (Section 3.4), which do not reduce to simple ratios and reflect movement relative to a common multi-period frontier; the stage gap supplements this evidence by showing whether a given period's overall productivity change was concentrated in lending or in loan-income generation, consistent with the research question of whether monetary policy normalization was associated with a change in the relative adjustment of the two stages, without asserting that the stage-level measures are themselves independent productivity indices.

3.6 Sample Construction and Merger Adjustment

The empirical sample covers FY2019 through FY2025. FY t is defined as April 1 of calendar year t through March 31 of $t + 1$. Accordingly, FY2023 runs from April 1, 2023 to March 31, 2024 and contains only the final 12 calendar days after the March 19, 2024 announcement; FY2024 and FY2025 are the first and second complete fiscal years after that announcement, respectively.

Bank-level deposits, loans, and loan interest income were transcribed from the annual financial statements and disclosure materials published by each sampled bank. We then construct a constant-perimeter balanced panel by retrospectively aggregating banks that subsequently merged. Before a legal merger, the deposits, loans, and loan interest income of the constituent banks are summed and assigned to the surviving entity. After the merger, the consolidated bank observation is used.

This procedure reduces the raw number of observed banks from 102 in FY2019 to a constant 96 banking entities in every fiscal year. The final panel therefore contains 672 bank-year observations and six annual productivity transitions.

Bank type and geographic region are fixed using the FY2025 classification of the surviving entity. This prevents changes in classification generated mechanically by mergers from appearing as changes in the composition of the regional or bank-type groups.

The six annual transitions are classified relative to the March 19, 2024 policy change. FY2019-FY2020, FY2020-FY2021, FY2021-FY2022, and FY2022-FY2023 provide the pre-normalization benchmark, while recognizing that the FY2023 endpoint contains 12 post-announcement days. FY2023-FY2024 compares an almost entirely pre-change fiscal year with the first complete post-change fiscal year. FY2024-FY2025 compares the first and second complete post-change fiscal years. This classification aligns observations with the event date but does not by itself identify a policy effect.

3.7 Statistical Comparison and Exploratory Regional Analysis

The DEA estimates are summarized across the 96 banking entities for each annual transition. Table 2 reports arithmetic means because the inferential analysis concerns bank-level additive differences between the paired stage-change indices. Because the Global MPI and both stage-change measures are positive ratios, geometric means are also calculated as a sensitivity check. Medians, sample standard deviations, minima, maxima, and interquartile ranges are used to assess distributional shape and the influence of extreme observations. The stage gap is an additive difference and is therefore summarized by the arithmetic mean and quantiles, not a geometric mean.

DEA estimates are deterministic conditional on the observed reference set. We therefore assess the sampling stability of the reported aggregates with a bank-cluster nonparametric bootstrap. Each of 10,000 replications resamples 96 analytical entities with replacement, carries each selected entity's complete FY2019–FY2025 trajectory into the resample, re-estimates the common multi-period frontier, and recalculates the Global MPI and stage measures. A fixed seed of 20260826 is used, and 95% percentile intervals are reported. Resampling complete trajectories preserves within-bank temporal dependence.

This procedure is a sampling-sensitivity analysis rather than the smoothed, bias-correcting frontier bootstrap of Simar and Wilson (1998). In the present scalar two-stage common-weight specification, the

multipliers cancel from consecutive-year stage ratios: lending-stage change reduces to the change in loans relative to deposits, and earnings-stage change reduces to the change in loan interest income relative to loans. Re-estimating the frontier can affect the Global MPI, but it does not alter an entity's stage-change ratios within a given resample. The bootstrap intervals therefore quantify the stability of the sample aggregates across banks; they are not bias-corrected confidence intervals for an underlying stochastic DEA frontier and do not address measurement error in the financial statements.

Two distinct comparisons should not be conflated. The primary comparison in this study is descriptive: because the panel covers the near-complete population of Japanese regional banking entities operating over FY2019–FY2025 (96 constant-perimeter entities consolidated from 102 pre-merger banks), differences between annual transitions or regions reported in Tables 2–5 describe the observed population directly and do not require a sampling model to be meaningful. The paired t-tests, Wilcoxon tests, Welch's ANOVA, and Kruskal–Wallis tests reported below, together with the bootstrap intervals described above, instead adopt a super-population inferential frame: they treat the observed banks and years as one realization drawn from a hypothetical larger population of possible bank trajectories, in order to gauge whether an observed difference is large relative to the variability such a process could plausibly generate. This inferential frame is a standard analytical convenience for interpreting the strength of observed patterns; it is not required to establish that the reported population-level differences occurred, and the hypothesis tests and bootstrap results throughout Sections [4](#) and [5](#) are presented as robustness and sensitivity checks on the descriptive population comparison, not as the primary basis for its substantive conclusions.

Since the two stage measures are obtained for the same banks, their differences are tested using paired t-tests. Wilcoxon signed-rank tests are also reported as a nonparametric check. Results for the four pre-normalization transitions are used as a benchmark for comparison with FY2023–FY2024 and FY2024–FY2025.

Regional differences are examined across seven geographic groups. Welch's ANOVA is used because group sizes and variances may differ, with Kruskal–Wallis tests reported as a nonparametric check. Differences between Regional Banks I and Regional Banks II are examined using Welch's two-sample tests.

To address multiplicity, raw p-values are retained and Benjamini–Hochberg false-discovery-rate adjusted q-values are calculated separately for four prespecified families: bank-level stage-gap contrasts, regional heterogeneity tests, bank-type comparisons, and the 20 exploratory correlation tests in Table 5. Statistical conclusions use $q < .05$. For Pearson correlations, conventional 95% confidence intervals are calculated using Fisher's z transformation. These intervals are reported to show the uncertainty created by the small regional sample; for pooled analyses they remain optimistic because the same seven regions are observed twice.

We also examine whether regional differences in the stage gap during FY2023–FY2024 and FY2024–FY2025 are associated with differences in loan-market concentration, deposit-market concentration,

regional loan growth, deposit growth, and changes in lending yield. Given the small number of regions (n = 7), these results are interpreted as exploratory associations rather than causal relationships.

Lending yield requires a different interpretation from the other regional variables. Stage 2 uses loans as its intermediate input and loan interest income as its output, so its score is closely related by construction to loan interest income per unit of loans. We retain the term earnings-stage change to describe movement in the DEA-based common-weight index, but it is not a price-adjusted measure of technical productivity. An increase can reflect the repricing of new or outstanding loans, differences in fixed- and variable-rate shares, loan maturity and credit composition, or other changes in income per unit of loans. The association with lending yield is therefore used only to describe the observed measure, not as an independent determinant or causal mechanism.

4. Results

4.1 Overall and Stage-Specific Efficiency and Productivity Change

Table 1 reports the mean overall and stage-specific efficiency scores for FY2019–FY2025. Lending-stage efficiency was higher than earnings-stage efficiency throughout the period. Overall efficiency declined from 0.476 in FY2019 to 0.455 in FY2021 and then increased gradually, reaching 0.507 in FY2025.

The two stages followed somewhat different paths. Mean lending-stage efficiency changed only moderately, from 0.550 in FY2019 to 0.549 in FY2023, before rising to 0.559 in FY2024 and 0.569 in FY2025. Earnings-stage efficiency fell from 0.344 in FY2019 to 0.318 in FY2021 and recovered to 0.330 by FY2023. It then increased to 0.345 in FY2024 and 0.403 in FY2025. The larger movement in FY2025 was therefore concentrated in the earnings stage.

Table 1
Overall and Stage-Specific Efficiency by Fiscal Year

Fiscal Year	Overall Efficiency	Lending Stage	Earnings Stage
FY2019	0.476	0.550	0.344
FY2020	0.458	0.532	0.324
FY2021	0.455	0.530	0.318
FY2022	0.462	0.543	0.319
FY2023	0.469	0.549	0.330
FY2024	0.481	0.559	0.345
FY2025	0.507	0.569	0.403
<i>Note. Values are means across the 96 banking entities. Overall and stage-specific efficiencies are calculated using the common optimal multipliers from the multi-period network model.</i>			

Table 2 reports arithmetic mean productivity-change measures for the six annual transitions from FY2019–FY2020 through FY2024–FY2025. Overall productivity changed relatively little during the four pre-normalization transitions. The arithmetic mean global Malmquist productivity index over this period was 0.997. Geometric-mean sensitivity estimates were very close to the reported arithmetic means in every transition.

Global productivity increased to 1.026 in FY2023–FY2024 and 1.057 in FY2024–FY2025. The stage-specific changes show more clearly how the two parts of the network differed. In FY2023–FY2024, mean lending-stage change was 1.022 and earnings-stage change was 1.053. In FY2024–FY2025, lending-stage change remained similar at 1.020, while earnings-stage change rose to 1.187.

The decomposition of global productivity also differed between the two transitions. In FY2023–FY2024, mean efficiency change was 1.135 and technical change was 0.904. In FY2024–FY2025, efficiency change was 1.002 and technical change was 1.055. The sources of the increase in global productivity were therefore different across the two periods.

Table 2
Overall and Stage-Specific Productivity Change

Transition	Period	Global MPI	Efficiency Change	Technical Change	Lending Stage	Earnings Stage	Stage Gap
2019→2020	Pre-normalization	0.963	0.953	1.011	0.966	0.941	-0.025
2020→2021	Pre-normalization	0.992	1.023	0.970	0.996	0.983	-0.013
2021→2022	Pre-normalization	1.018	1.000	1.017	1.026	1.009	-0.017
2022→2023	Pre-normalization	1.015	0.928	1.093	1.011	1.037	0.026
2023→2024	First post-change comparison	1.026	1.135	0.904	1.022	1.053	0.032
2024→2025	Second post-change comparison	1.057	1.002	1.055	1.020	1.187	0.167

Note. Entries are arithmetic means across the 96 banking entities. Global MPI > 1 indicates productivity improvement. Lending- and earnings-stage changes are calculated for each entity from consecutive-year stage efficiencies. Stage Gap = Earnings-stage common-weight change – Lending-stage common-weight change. The pre-normalization arithmetic mean Global MPI was 0.997 and the mean Stage Gap was – 0.007. Geometric means are reported in the accompanying sensitivity discussion because they are defined for the positive ratio indices but not for the additive stage gap.

Distributional sensitivity for FY2024–FY2025 supports the same interpretation. For the Global MPI, the arithmetic mean was 1.057, the geometric mean 1.056, and the median 1.056 (SD = 0.029; interquartile range 1.040–1.071; range 0.985–1.178). For lending-stage change, the corresponding values were 1.020, 1.019, and 1.014 (SD = 0.042; interquartile range 0.999–1.038; range 0.828–1.197). For earnings-stage change, they were 1.187, 1.182, and 1.170 (SD = 0.118; interquartile range 1.132–1.230; range 0.989–1.887). Although the earnings-stage distribution includes a high maximum, its geometric mean and median remain well above one and close to the arithmetic mean; the second-year increase is therefore not attributable solely to a few extreme observations.

The bank-cluster bootstrap gives the same qualitative conclusion for FY2024–FY2025. The mean Global MPI was 1.057 (95% bootstrap interval, 1.051–1.077), mean lending-stage change was 1.020 (1.012–1.028), mean earnings-stage change was 1.187 (1.165–1.213), and the mean stage gap was 0.167 (0.142–0.195). All four intervals remain above their no-change benchmark for the relevant measure.

4.2 Widening Stage Gap after Monetary Policy Normalization

Figure 1 shows the mean stage gap for each annual transition. The gaps for the four pre-normalization transitions were -0.0255 , -0.0134 , -0.0169 , and 0.0264 , with an average of -0.0074 . There was thus little difference between lending-stage and loan-income-stage performance change before normalization.

The gap rose to 0.0319 in FY2023–FY2024 and to 0.1667 in FY2024–FY2025. In FY2024–FY2025, mean lending-stage change was 1.0200 and mean earnings-stage change was 1.1867 . The wider gap was therefore driven mainly by the increase in the earnings-stage common-weight index rather than by deterioration in the lending-stage index. This is a result about relative measured performance; the index does not by itself distinguish repricing from a change in production technology.

Note

The 95% bank-cluster bootstrap interval is shown for FY2024-FY2025, the transition for which the bootstrap interval was estimated. q is the Benjamini-Hochberg-adjusted value for the paired stage-gap comparison.

For FY2024–FY2025, the paired mean difference between earnings- and lending-stage change was 0.1667 , with a parametric 95% confidence interval of 0.1399 to 0.1935 and a bank-cluster bootstrap interval of 0.142 to 0.195 . The difference was significant in both the paired t-test, $t(95) = 12.34$, $p < .001$, and the Wilcoxon signed-rank test ($p < .001$). These results remain significant after Benjamini–Hochberg adjustment within the stage-gap comparison family ($q < .005$). Earnings-stage change was higher than lending-stage change for 91 of the 96 banks (94.8%); the bootstrap interval for this proportion was 89.6% to 99.0%.

Compared with each bank's average gap over the four pre-normalization transitions, the mean gap was 0.0392 higher in FY2023–FY2024 ($p < .001$) and 0.1740 higher in FY2024–FY2025 ($p < .001$). The gap also increased by 0.1348 from FY2023–FY2024 to FY2024–FY2025 ($p < .001$). All three contrasts remain significant after Benjamini–Hochberg adjustment ($q < .005$).

The difference between the two stages was small before normalization and widened most clearly in FY2024–FY2025. This reflected the stronger increase in the earnings-stage common-weight index, while the lending-stage index remained slightly above one. Given the Stage 2 variables, the increase should not be read as evidence of technical productivity improvement alone.

4.3 Regional Heterogeneity

Table 3 reports the stage gap for the seven geographic regions. Regional differences were relatively small in FY2023–FY2024. Mean gaps ranged from -0.001 in Tokai to 0.062 in Koshinetsu-Hokuriku. Neither Welch's ANOVA ($p = .541$, $q = .541$) nor the Kruskal–Wallis test ($p = .320$, $q = .384$) indicates significant regional differences.

The differences were much larger in FY2024–FY2025. The mean stage gap was 0.235 in Tokai, 0.227 in Kinki, 0.220 in Hokkaido-Tohoku, and 0.192 in Koshinetsu-Hokuriku. The corresponding gaps were 0.154 in Kanto, 0.136 in Kyushu-Okinawa, and 0.081 in Chugoku-Shikoku. Regional differences remain significant after adjustment in both Welch's ANOVA and the Kruskal–Wallis test (both $p < .001$; $q < .003$).

Table 3
Regional Heterogeneity in the Stage Gap

Region	N	Pre-normalization	FY2023–FY2024	FY2024–FY2025	Change from FY2023–FY2024	FY2024–FY2025 – Pre
Hokkaido-Tohoku	16	–0.021	0.019	0.220	0.201	0.241
Kanto	15	–0.004	0.031	0.154	0.122	0.157
Koshinetsu-Hokuriku	10	–0.009	0.062	0.192	0.131	0.201
Tokai	10	–0.012	–0.001	0.235	0.236	0.247
Kinki	8	–0.024	0.061	0.227	0.166	0.251
Chugoku-Shikoku	17	0.019	0.038	0.081	0.043	0.061
Kyushu-Okinawa	20	–0.012	0.027	0.136	0.109	0.148
Note. Values are mean stage gaps. Raw p-values and Benjamini–Hochberg q-values for the six regional heterogeneity tests are: FY2023–FY2024, Welch $p = .541$, $q = .541$ and Kruskal–Wallis $p = .320$, $q = .384$; FY2024–FY2025, both $p < .001$ and $q < .003$; change between the two transitions, Welch $p = .006$, $q = .009$ and Kruskal–Wallis $p = .001$, $q = .002$.						

The increase in the stage gap from FY2023–FY2024 to FY2024–FY2025 also varied across regions. The increase was largest in Tokai (0.236) and Hokkaido-Tohoku (0.201), and smallest in Chugoku-Shikoku (0.043). The regional differences remain significant under both Welch's ANOVA ($p = .006$, $q = .009$) and the Kruskal–Wallis test ($p = .001$, $q = .002$). These regional patterns are summarized in Fig. 2.

Note

q-values are Benjamini-Hochberg adjusted. The displayed q-values use Welch's ANOVA; Kruskal-Wallis conclusions are the same (FY2023-FY2024, $q = .384$; FY2024-FY2025, $q < .003$; change, $q = .002$).

4.4 Regional Banks I and Regional Banks II

Table 4 compares the stage gap between Regional Banks I and Regional Banks II. Before normalization, the mean stage gap was 0.0005 for Regional Banks I and – 0.0210 for Regional Banks II. The raw

difference was significant and remains just below the 5% threshold after adjustment across the four bank-type comparisons ($p = .011$, $q = .044$).

The difference between the two groups was much smaller in FY2023–FY2024 and FY2024–FY2025. In FY2023–FY2024, the mean gap was 0.0323 for Regional Banks I and 0.0311 for Regional Banks II ($p = .947$, $q = .947$). In FY2024–FY2025, the corresponding values were 0.1676 and 0.1650 ($p = .930$, $q = .947$).

Table 4
Regional Banks I versus Regional Banks II

Bank Type	N	Pre-normalization	FY2023–FY2024	FY2024–FY2025	FY2024–FY2025 – Pre
Regional Banks I	61	0.000	0.032	0.168	0.167
Regional Banks II	35	–0.021	0.031	0.165	0.186
Note. Values are mean stage gaps. Welch's two-sample tests with Benjamini–Hochberg adjustment across the four bank-type comparisons: pre-normalization, $p = .011$, $q = .044$; FY2023–FY2024, $p = .947$, $q = .947$; FY2024–FY2025, $p = .930$, $q = .947$; change from the pre-normalization mean to FY2024–FY2025, $p = .575$, $q = .947$.					

The increase from the pre-normalization mean to FY2024–FY2025 was 0.1671 for Regional Banks I and 0.1860 for Regional Banks II. This difference was not statistically significant ($p = .575$).

4.5 Exploratory Regional Analysis

Table 5 examines the association between the stage gap and regional banking conditions. The variables include changes in lending yield, loan growth, deposit growth, loan-market concentration, and deposit-market concentration. The analysis is based on seven regions and is exploratory.

When FY2023–FY2024 and FY2024–FY2025 are pooled, the stage gap is closely associated with the change in regional lending yield. The Pearson correlation is 0.960 (conventional 95% CI, 0.875 to 0.988; $p < .001$, $q < .010$), and the Spearman correlation is 0.943 ($p < .001$, $q < .010$). These pooled results use 14 region-year observations but only seven distinct regions, so the confidence interval and adjusted p -values do not remove the dependence created by repeated regional observations.

The separate-year estimates are much less precise. In FY2023–FY2024, the stage gap is correlated with regional loan growth ($r = .871$; 95% CI, .343 to .981; $p = .011$, $q = .073$) and lending-yield change ($r = .845$; 95% CI, .253 to .977; $p = .017$, $q = .085$). In FY2024–FY2025, the correlation with loan growth is negative ($r = -.424$; 95% CI, $-.892$ to .483; $p = .344$, $q = .415$), while the correlation with lending-yield change remains positive ($r = .815$; 95% CI, .160 to .972; $p = .025$, $q = .097$). None of these separate-year correlations remains significant after adjustment.

Loan- and deposit-market concentration are also positively correlated with the stage gap in FY2024–FY2025, but the uncertainty is substantial: loan HHI, $r = .687$ (95% CI, $-.137$ to $.949$; $p = .088$, $q = .180$), and deposit HHI, $r = .704$ (95% CI, $-.104$ to $.952$; $p = .078$, $q = .180$). Neither result is statistically significant.

Table 5
Exploratory Correlations between Regional Conditions and the Stage Gap

Variable	Pooled Pearson r	p	Pooled Spearman ρ	p	FY2023–FY2024 Pearson r	p	FY2024–FY2025 Pearson r	p
Lending yield change	0.960	< .001	0.943	< .001	0.845	0.017	0.815	0.025
Loan growth	0.389	0.169	0.499	0.069	0.871	0.011	–0.424	0.344
Deposit growth	0.470	0.090	0.582	0.029	0.587	0.166	–0.000	1.000
Loan HHI	0.272	0.347	0.235	0.418	0.355	0.435	0.687	0.088
Deposit HHI	0.286	0.322	0.284	0.326	0.417	0.353	0.704	0.078
Note. Raw p-values are shown in the table. Benjamini–Hochberg adjustment was applied across all 20 correlation tests. Only the pooled lending-yield Pearson and Spearman associations have $q < .05$ (both $q < .010$); all separate-year correlations have $q \geq .073$. The pooled analysis contains 14 region-year observations but only seven distinct regions, and the separate-year analyses each have $n = 7$. Conventional Fisher-z confidence intervals are therefore wide and the pooled inference is potentially anti-conservative. Lending yield is also mechanically related to the Stage 2 production structure. All correlations are treated as exploratory rather than confirmatory.								

Lending yield is calculated from loan interest income and loans, which also enter the earnings stage. Its correlation with the stage gap therefore partly reflects the structure of the Stage 2 measure itself. We use this result to help interpret the regional pattern in measured loan-income-stage performance, not as evidence of a causal relationship or technical productivity improvement.

A large language model (ChatGPT, OpenAI) was used during preparation of this manuscript to assist with language editing, the structural organization of the methodology and discussion, and the presentation of the robustness checks described above (Section 3.7). It was not used to generate, select, or verify the underlying bank-level data, and it was not used to derive the DEA and Malmquist formulas in Sections 3.2–3.5, which follow Kourtzidis et al. (2021) and Pastor and Lovell (2005). All analytical choices, computations, and reported results were independently verified by the author, who takes full responsibility for the content of this manuscript.

5. Discussion

5.1 Monetary Policy Normalization and the Rebalancing of Bank Intermediation

Overall measured productivity increased after monetary policy normalization, but the two stage-specific common-weight indices moved unevenly. Lending- and earnings-stage changes had been relatively close before normalization. The difference remained small in FY2023–FY2024 and became much larger in FY2024–FY2025. This finding describes a reallocation of measured performance within the network and does not, by itself, identify a change in underlying production technology.

The two stages capture different parts of the banking process. The lending stage concerns the transformation of deposits into loans, while the earnings stage concerns loan interest income generated from those loans. In FY2024–FY2025, the lending-stage change index remained slightly above one, whereas the earnings-stage index rose substantially. Because the latter is closely related to interest income per unit of loans, its movement can reflect price and portfolio effects that are conceptually distinct from technical productivity.

The widening stage gap therefore came mainly from the earnings side rather than from weaker lending performance. We describe this pattern as an earnings-led rebalancing of measured intermediation performance. This wording is deliberately narrower than a claim of technical productivity improvement: banks may have generated more loan interest income per unit of loans because existing loans repriced at different speeds, because new lending carried higher rates, or because loan portfolios differed in maturity, rate type, or credit composition.

The distinction is important when assessing the effects of monetary policy normalization. The DEA result establishes that the two stage indices diverged, but it does not show that banks produced more real intermediation services from the same resources. Without transaction-level rates, repricing schedules, and loan-composition data, the price, mix, and operational components of the earnings-stage increase cannot be separated.

5.2 Why Did the Gap Widen More in the Second Year?

One possible explanation for the larger second-year gap is the time required for changes in interest rates to pass through loan portfolios. Monetary policy changed in March 2024, but rates on outstanding loans did not change immediately. New loans can reflect higher market rates sooner, while existing loans are repriced according to their contractual terms. This mechanism is economically plausible, but the present data do not identify it separately from changes in loan composition or other bank-specific factors.

The results show a clear difference between the first and second post-change comparisons. The stage gap was positive but relatively small in FY2023–FY2024 and increased substantially in FY2024–FY2025. The productivity decomposition also differed. Efficiency change was relatively large in the first comparison, while technical change was below one. In the second comparison, efficiency change was close to one and technical change was above one.

The regional results show a similar change. Loan growth was positively related to the stage gap in the first year, but not in the second. Lending-yield change, however, was strongly related to the gap in both years. By the second year, the larger stage gap was therefore more closely associated with changes in income from lending than with loan growth.

There is an important limitation to this interpretation. Lending yield is calculated from the same loans and loan interest income that define Stage 2, so its relationship with the earnings-stage index and the stage gap is partly mechanical. The correlation cannot establish repricing as the cause of the widening gap. The defensible conclusion is narrower: measured loan-income performance rose relative to lending performance in the second year, while the underlying contributions of prices, portfolio mix, risk, and operational productivity remain unidentified.

5.3 Regional Heterogeneity after Normalization

Regional differences were small in the first post-change comparison and much larger in the second. In FY2024–FY2025, the stage gap was largest in Tokai, Kinki, Hokkaido-Tohoku, and Koshinetsu-Hokuriku, while Chugoku-Shikoku showed a much smaller increase. The regional differences were statistically significant in the second year, but not in the first.

The emergence of these differences in the second year is important. Monetary policy normalization applied to all regions at the same time, but the stage gap did not develop in the same way across regional banking markets. The first-year results were relatively similar across regions. Much of the regional variation appeared later.

Market concentration may account for some of this variation, although the evidence is weak. Both loan and deposit concentration were positively correlated with the second-year stage gap, but their 95% confidence intervals include zero and their multiplicity-adjusted q-values equal .180. With only seven regions, there is not enough evidence to conclude that concentration explains the regional differences.

What can be said from the results is that the second-year widening was uneven across regions. National averages conceal this variation. Explaining why some regions experienced a much larger earnings-lending gap than others requires evidence beyond the variables examined here.

5.4 Bank Type and Geographic Differences

Regional Banks I and Regional Banks II differed before normalization, but this difference disappeared after normalization. Their mean stage gaps were almost the same in both post-change comparisons. The increase from the pre-normalization period also did not differ significantly between the two groups.

Kourtzidis et al. (2021) found differences in inefficiency between Regional Banks I and Regional Banks II for the period from 2002 to 2017. Our results are different in focus. We examine changes in the two stages around monetary policy normalization rather than differences in inefficiency between the two bank types. In this setting, the distinction between Regional Banks I and II does not explain the post-normalization stage gap.

The regional results are different. There was little regional variation in the first post-change comparison, but clear differences appeared in the second. No similar difference appeared between Regional Banks I and Regional Banks II.

This does not mean that geography explains the stage gap. The regional analysis is based on only seven groups, and the variables examined in Section 4.5 provide limited evidence on the source of the differences. What the results show is simpler: after normalization, differences in the stage gap appeared across regions, but not between Regional Banks I and Regional Banks II.

5.5 Implications for Measured Bank Performance during Policy Normalization

Looking only at overall productivity would miss an important part of the post-normalization change. Overall productivity increased in both post-change comparisons, but the two stage-specific indices did not move in the same way. The difference was especially large in FY2024–FY2025, when the earnings-stage index increased much more than the lending-stage index. This divergence is informative about where the measured change occurred, even though its price and quantity components cannot be separated.

The contrast between the two post-change transitions shows why adjustment should be evaluated over more than one annual window. The divergence was modest in the initial comparison and became substantially more pronounced in the following year; an analysis confined to the immediate post-change transition would therefore miss the delayed pattern.

The results also require care in interpreting stronger bank performance after normalization. A higher earnings-stage index indicates more loan interest income relative to loans under the common multiplier structure; it is not direct evidence that banks became technically more productive. In FY2024–FY2025, the lending-stage index remained close to its first-year level while the earnings-stage index rose sharply. Whether this reflected faster repricing, loan composition, risk compensation, operational improvement, or a combination of these factors is beyond what the available data can determine.

Finally, the size of the stage gap differed across regions. These differences were small in the first post-change comparison and much larger in the second. National results therefore describe the general direction of change, but not the size of the adjustment experienced across regional banking markets.

5.6 Limitations and Further Research

The analysis has several limitations. Only two years after monetary policy normalization are included. This is a short period, especially because the results differ considerably between the first and second years. It is too early to know whether the increase in the earnings-stage index observed in FY2024–FY2025 will continue.

The analysis also does not establish that monetary policy normalization caused the observed changes. All Japanese regional banks faced the same national policy shift, leaving no contemporaneous untreated domestic group for a difference-in-differences or comparable counterfactual design. The exercise should therefore be read as a descriptive event-aligned before-after comparison. March 19, 2024 is a timing marker rather than an isolated exogenous treatment. Concurrent developments-including inflation and nominal activity, wage growth and changes in household and corporate cash flows, subsequent Bank of Japan actions and movements in short- and long-term market rates, credit demand and borrower risk, regional business-cycle and demographic conditions, equity-market valuations and bank capital conditions, and deposit competition and funding costs-could also affect deposits, loans, and loan interest income. With the present three-variable production structure and only two complete post-change fiscal years, these channels cannot be separately controlled or attributed.

The DEA model is limited to deposits, loans, and loan interest income. This specification separates lending from loan-income generation, but Stage 2 is consequently close to a lending-yield measure and is not adjusted for changes in interest rates, credit risk, loan maturity, fixed- versus variable-rate composition, or other portfolio characteristics. It also excludes securities and fee income. The results therefore concern measured performance within this narrow loan-intermediation structure rather than technical productivity or whole-bank performance.

The bootstrap sensitivity analysis shows that the main FY2024–FY2025 aggregates are stable to resampling banks and re-estimating the observed frontier. It does not convert the deterministic DEA model into a stochastic production model. In particular, it does not correct financial-statement measurement error, specification error, or omitted inputs and outputs, and it is not a substitute for a model-specific Simar–Wilson smoothed bootstrap adapted to the multi-period relational network structure.

The regional results need particular caution. There are only seven regions, which leaves little scope for statistical analysis beyond the reported comparisons and correlations. Moreover, loans and loan interest income enter both lending yield and Stage 2. The strong correlation with the stage gap is therefore structurally induced in part and is not independent evidence of repricing or productivity improvement.

A longer post-normalization period will show whether the difference between lending and earnings persists. Bank-level data on contractual reset dates, new-loan and outstanding-loan rates, fixed- and variable-rate shares, maturity, borrower risk, and loan composition would allow future research to separate repricing and portfolio effects from changes in operational productivity.

6. Conclusion

This study applies a multi-period additive relational network DEA framework, with the Global Malmquist productivity index and its efficiency-change/technical-change decomposition as its primary DEA-derived measure of productivity change, to a balanced panel of Japanese regional banks spanning the BOJ's 2024 policy normalization. Its contribution is an event-aligned assessment of how measured

intermediation performance changed at the whole-network level and, through an auxiliary ratio decomposition consistent with the network structure, of where within the connected lending-and-earnings process that change was concentrated, rather than a causal estimate of the policy's effect.

The evidence shows that overall productivity, measured by the Global MPI, increased after normalization, and that the two stage-level ratios—the loan-to-deposit ratio and loan interest income per unit of loans—moved similarly before normalization but diverged afterward, especially in the later comparison, as the loan-yield ratio rose more strongly than the loan-to-deposit ratio. Regional heterogeneity also became more visible, while the two institutional bank categories remained broadly similar. These patterns show that an aggregate productivity index can conceal uneven adjustment within the intermediation network, even when the stage-level diagnostic used to reveal that unevenness reduces to ordinary financial ratios rather than to independent efficiency constructs.

The interpretation is deliberately narrow, and so is the methodological claim. Because the second-stage ratio relates loan interest income to loans, its movement may reflect repricing speed, portfolio composition, and risk compensation as well as operational performance; it is not a price-adjusted measure of technical productivity. Because the ratio reduction demonstrated in Section 3.5 follows directly from the single-input/single-output stage specification, it will recur in any two-stage network DEA model built on this structure; the contribution offered here is therefore a transparent, common-weight-consistent way of organizing such ratio diagnostics alongside genuine frontier-based productivity measures, generalizable to other single-input/single-output network settings, rather than a new stage-level efficiency index. Longer post-normalization panels, bank-level information on contractual reset dates, rate structures, maturity, and borrower risk, and network specifications with additional inputs or outputs are needed to distinguish repricing and portfolio channels from operational productivity and to recover genuine multi-dimensional stage efficiency scores.

Declarations

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Authorship Contributions

The author was solely responsible for study conception and design, data collection and construction of the analytical panel, formal analysis, interpretation of results, and drafting and revising the manuscript. The author read and approved the final manuscript.

Declaration of Competing Interest

The author declares that there are no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data Availability

The underlying observations were collected from annual financial statements and disclosure materials published by the sampled Japanese regional banks, including materials available through the banks' investor-relations and disclosure websites. These publicly issued documents are the primary sources. The authors used them to construct a harmonized bank-year panel containing deposits, loans, and loan interest income for 96 constant-perimeter banking entities over FY2019-FY2025 (672 bank-year observations), together with a merger-adjustment crosswalk. The compiled analytical panel and crosswalk, rather than the underlying public documents themselves, are available from the corresponding author upon reasonable request. No proprietary database is required to reconstruct the reported input-output panel from the identified bank disclosures.

Ethics Approval and Consent to Participate

This study uses only publicly disclosed, non-confidential financial-statement and disclosure data for regional banking entities. It does not involve human participants, personal data, or animal subjects. Ethics approval and consent to participate were therefore not required.

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work, the author used ChatGPT (OpenAI) to assist with language editing, structural revision, and the presentation of robustness information; the tool was not used as a source of bank data. Analytical choices, reported computations, source verification, and the final manuscript remain the author's sole responsibility.

References

1. Bank of Japan (2024a) Changes in the monetary policy framework. Bank of Japan, Tokyo
2. Bank of Japan (2024b) Financial system report: October 2024. Bank of Japan, Tokyo
3. Banker RD, Charnes A, Cooper WW (1984) Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Manag Sci* 30(9):1078–1092. <https://doi.org/10.1287/mnsc.30.9.1078>
4. Borio C, Gambacorta L, Hofmann B (2017) The influence of monetary policy on bank profitability. *Int Finance* 20(1):48–63. <https://doi.org/10.1111/infi.12104>
5. Charnes A, Cooper WW, Rhodes E (1978) Measuring the efficiency of decision making units. *Eur J Oper Res* 2(6):429–444. [https://doi.org/10.1016/0377-2217\(78\)90138-8](https://doi.org/10.1016/0377-2217(78)90138-8)
6. Chen Y, Cook WD, Li N, Zhu J (2009) Additive efficiency decomposition in two-stage DEA. *Eur J Oper Res* 196(3):1170–1176. <https://doi.org/10.1016/j.ejor.2008.05.011>
7. Chen Y, Cook WD, Zhu J (2010) Deriving the DEA frontier for two-stage processes. *Eur J Oper Res* 202(1):138–142. <https://doi.org/10.1016/j.ejor.2009.05.012>

8. Chen Y, Zhu J (2004) Measuring information technology's indirect impact on firm performance. *Inf Technol Manag* 5(1–2):9–22. <https://doi.org/10.1023/B:ITEM.0000008075.43543.97>
9. Claessens S, Coleman N, Donnelly M (2018) Low-for-long interest rates and banks' interest margins and profitability: cross-country evidence. *J Financ Intermed* 35:1–16. <https://doi.org/10.1016/j.jfi.2017.05.004>
10. Färe R, Grosskopf S, Norris M, Zhang Z (1994) Productivity growth, technical progress, and efficiency change in industrialized countries. *Am Econ Rev* 84(1):66–83
11. Fukuyama H (1995) Measuring efficiency and productivity growth in Japanese banking: a nonparametric frontier approach. *Appl Financ Econ* 5(2):95–107
12. Fukuyama H, Matousek R (2017) Modelling bank performance: a network DEA approach. *Eur J Oper Res* 259(2):721–732. <https://doi.org/10.1016/j.ejor.2016.10.044>
13. Fukuyama H, Matousek R, Tzeremes NG (2022) Bank production with nonperforming loans: a minimum distance directional slack inefficiency approach. *Omega* 113:102706. <https://doi.org/10.1016/j.omega.2022.102706>
14. Hoshi T, Kashyap AK (2004) Japan's financial crisis and economic stagnation. *J Econ Perspect* 18(1):3–26. <https://doi.org/10.1257/089533004773563412>
15. Kourtzidis S, Matousek R, Tzeremes NG (2021) Modelling a multi-period production process: evidence from the Japanese regional banks. *Eur J Oper Res* 294(1):327–339. <https://doi.org/10.1016/j.ejor.2021.01.036>
16. Pastor JT, Lovell CAK (2005) A global Malmquist productivity index. *Econ Lett* 88(2):266–271. <https://doi.org/10.1016/j.econlet.2005.02.013>
17. Simar L, Wilson PW (1998) Sensitivity analysis of efficiency scores: how to bootstrap in nonparametric frontier models. *Manag Sci* 44(1):49–61. <https://doi.org/10.1287/mnsc.44.1.49>
18. Shi Y, Charles V, Zhu J (2025) Bank financial sustainability evaluation: data envelopment analysis with random forest and Shapley additive explanations. *Eur J Oper Res* 321(2):614–630. <https://doi.org/10.1016/j.ejor.2024.09.030>
19. Tone K, Tsutsui M (2009) Network DEA: a slacks-based measure approach. *Eur J Oper Res* 197(1):243–252. <https://doi.org/10.1016/j.ejor.2008.05.027>

Figures

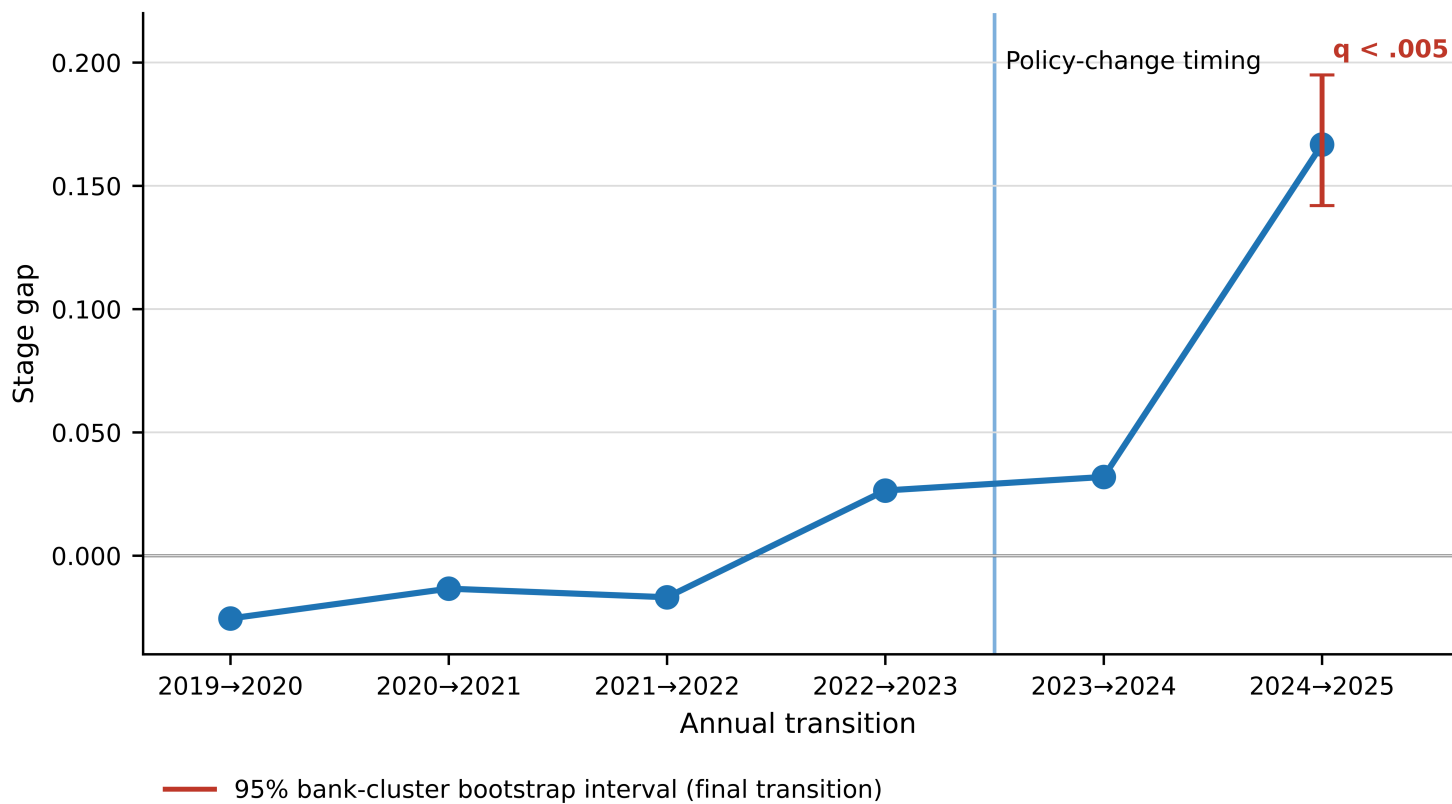


Figure 1

Stage Gap across Annual Transitions

Note. The 95% bank-cluster bootstrap interval is shown for FY2024-FY2025, the transition for which the bootstrap interval was estimated. q is the Benjamini-Hochberg-adjusted value for the paired stage-gap comparison.

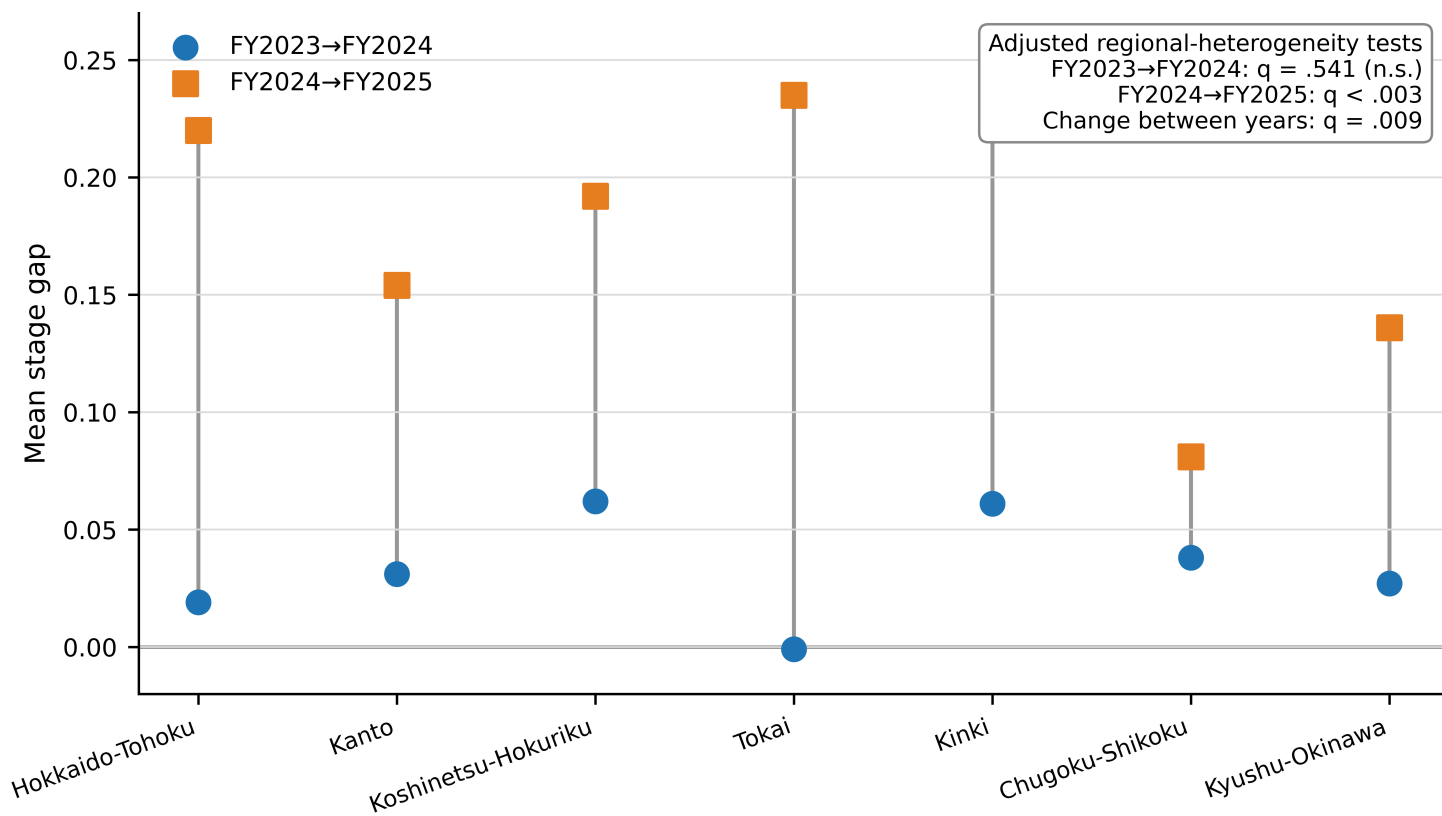


Figure 2

Regional Stage Gap in FY2023–FY2024 and FY2024–FY2025

Note. q-values are Benjamini-Hochberg adjusted. The displayed q-values use Welch's ANOVA; Kruskal-Wallis conclusions are the same (FY2023-FY2024, $q = .384$; FY2024-FY2025, $q < .003$; change, $q = .002$).