Banking Reform, Credit Allocation, and Local Real Effects: Evidence from India*

Yogeshwar Bharat† Suleyman Faruk Gozen ‡
David Hong § Mehmet Furkan Karaca ¶

Abstract

We leverage the Reserve Bank of India's 2006 Bank Authorization Policy as a quasi-natural experiment to study effects on credit markets, capital misallocation, and firm outcomes. We find asymmetric responses: private-sector bank branches expanded by 16.3% in underbanked districts relative to banked districts, while public-sector banks showed no systematic expansion. The resulting private-sector lending reduced the marginal revenue product of capital (MRPK) of ex-ante high-MRPK firms by about 60%, lowering capital misallocation. However, this decline did not raise firm sales or value added. We highlight the efficacy of financial reforms in alleviating misallocation under mixed-ownership banking environments in developing economies.

Keywords: Government Policies, Bank Branch Expansion, Local Credit Market, Capital Misallocation, Firm Outcomes, Financial Intermediation in Developing Economies, Bank Ownership

JEL codes: E22, E44, E51, G21, G28, O16

---

*We thank Arpad Abraham, David Afikuyomi, Chloe Ahn, Sekyu Choi, Pawel Doligalski, Sizhe Hong, Chun Kuang, Gizem Kutlu, Tai-Wei Hu, César Garro-Marín, Iacopo Morchio, Hakki Yazici as well as workshop/conference participants at the 10th International Conference on Applied Theory, Macro and Empirical Finance (AMEF), 2025 Annual Conference of the Scottish Economic Society (SES), and Bristol Macro Brownbag for their very helpful and constructive comments. Computer programs to replicate our findings are available from the authors. All remaining errors are our own.

†Department of Economics, Shiv Nadar Institution of Eminence. Email: yogeshwar.bharat@snu.edu.in.

‡School of Economics, University of Bristol. Email: suleyman.gozen@bristol.ac.uk.

§School of Economics, The University of Edinburgh. Email: dhong@ed.ac.uk.

¶Essex Business School, University of Essex. Email: m.f.karaca@essex.ac.uk.

1 Introduction

Financial intermediaries are essential for facilitating credit and driving resource allocation. The consequences of less well-developed financial markets exacerbating capital misallocation and ultimately dampening economic productivity have been well-documented (see Gopinath et al., 2017; Restuccia and Rogerson, 2008, 2017, among others). Thus, regulators in developing economies are keen on enacting financial reforms to rectify these consequences. However, the real effects of banking reforms at the local level remain a less understood topic. In this study, we examine the efficacy of one such reform in India in 2006 and show a nuanced result that the ex-ante composition of operating banks by bank ownership plays a major factor in determining a policy's success in local markets.

The Indian credit market is an ideal laboratory due to the heterogeneity in bank ownership. Public-sector banks (government-owned banks, henceforth PSBs) account for two-thirds of total deposits and issue 70% of total loans, with the remainder operated by Private-sector banks (hereafter PVBs). The empirical evidence on the usefulness of PSBs has been mixed. Altunbas et al. (2001) find no efficiency advantage for German PVBs relative to PSBs, while Sapienza (2004) and Micco et al. (2007) report varied positive impacts of PSBs. Within the Indian context, Cole (2009) find that PSBs tend to increase agricultural lending to districts with close races during election years. Banerjee et al. (2004) and Kumar (2020) reinforce this, finding that PSBs allocate more credit to agriculture and rural areas, often at the expense of manufacturing firms. These institutions also frequently prioritize social objectives, such as financial inclusion and lending to specific sectors at the expense of efficiency (see Carvalho, 2014; Coleman and Feler, 2015). This dynamic interplay between PSBs and PVBs enables us to examine the effectiveness of financial reforms in developing economies through the bank credit channel.

We leverage the Bank Authorization Policy (BAP) implemented by the Reserve Bank of India (RBI) in 2006 as a quasi-natural experiment to capture exogenous variation and provide a causal analysis. This policy standardized the branch expansion process for Indian banks and classified districts as banked or underbanked based on whether their bank branch density was above or below the national average. Thus, the RBI's financial inclusion objective would elicit bank branch expansions to occur in these underbanked districts. Several papers also exploit the same quasi-natural experiment. Young (2017) find that BAP led to an increase in local GDP growth in underbanked districts. While Kulkarni et al. (2023) find that credit-constrained establishments increased capital expenditures closer in distance to a bank branch. Rather than focusing on aggregate growth outcomes (Young, 2017) or establishment-level investment responses (Kulkarni et al., 2023), we analyze the Bank Authorization Policy through the institutional lens of bank ownership, examining its impact on the dispersion of marginal products of capital across firms within districts and on firm-level outcomes.

This paper addresses key questions about the impact of the policy: What is the effect of the policy on branch expansions in underbanked districts, and how does bank ownership play a role? How does the policy affect capital misallocation and related firm outcomes among manufacturing firms at the district level? Finally, what mechanisms link bank lending to capital misallocation and firm outcomes? We address these questions using a difference-in-differences methodology, providing causal evidence on three main dimensions: (i) the policy's impact on bank branch expansion by bank ownership type, (ii) its effects on district-level capital misallocation and related firm outcomes, and (iii) the mechanism linking these outcomes through the bank-to-firm lending channel.

We draw on three complementary datasets: (i) RBI's Basic Statistical Returns, which record branch openings and industrial credit growth by ownership at the district level; (ii) firm-level financial statements from the CMIE Prowess database, which we use to measure firm-level outcomes and calculate marginal revenue product of capital (MRPK) as a measure of capital misallocation; and (iii) matched firm–bank loan data from the Ministry of Corporate Affairs, which link loan issuance to firm characteristics and bank ownership. Following Sraer and Thesmar (2023b,a), we treat MRPK as a sufficient statistic for misallocation, consistent with models where financial frictions distort capital allocation (Restuccia and Rogerson, 2008; Hsieh and Klenow, 2009). We expect the policy change to alleviate financial constraints through the bank credit channel.

We provide evidence on the policy's impact from several complementary angles, tracing its effects on bank branch expansion, bank-to-firm credit allocation by bank ownership type, capital misallocation, and key firm outcomes. First, we show that the 2006 BAP increased bank branch expansion in underbanked districts, but only through private-sector banks (PVBs). Our difference-in-differences estimates indicate that PVBs expanded branches by about 16 percent in underbanked districts relative to banked ones after the reform. At the same time, public-sector banks (PSBs)—despite their dominant market share—showed no response. This asymmetric response highlights the central role of ownership structure: branch expansion mandates can alter private banks' incentives to enter underserved areas, but do little to reshape the behavior of public banks. The policy, therefore, shifted the composition of lending in underbanked districts decisively toward private-sector banks.

Second, we find that the additional credit from PVB branch expansion was disproportionately absorbed by ex-ante high-MRPK firms. Using matched bank–firm loan data, we document that PVBs selectively extended new loans to firms with above-average MRPK in the pre-policy period. These borrowers tended to be larger, cash-rich, less leveraged, and more mature—firms with stronger balance sheets and lower default risk. In contrast, ex-ante low-MRPK firms in the same districts saw little to no increase in borrowing from either PVBs or PSBs.

Third, this targeted reallocation of credit reduced capital misallocation. Firms with higher pre-policy MRPK in underbanked districts expanded their capital base significantly after the reform, while lower-MRPK firms did not. Yet, the capital expansion did not immediately translate into higher sales. As a result, the dispersion in MRPK across firms declined by nearly 60 percent, consistent with more efficient capital allocation. The findings highlight an important mechanism: the policy reduced misallocation through improved credit allocation by PVBs in underbanked districts, but the gains were short-lived. They did not translate into sales growth in the short run.

Fourth, complementary firm outcomes reinforce this interpretation. High-MRPK firms in underbanked districts increased capital per worker and improved their profit margins following the reform, suggesting that new credit was used to deepen capital intensity and strengthen profitability. However, we find no significant effect on value added. Low-MRPK firms also show no meaningful changes across any dimension—credit access, capital deepening, or profitability—providing evidence on the asymmetric benefits of the policy. This disconnect between profitability gains and stagnant value-added growth highlights the limits of branch expansion as a growth policy: while it improved allocative efficiency, capital deepening, and profitability for ex-ante high-MRPK firms, it failed to generate sustained increases in value added.

Taken together, our evidence shows how the 2006 BAP unfolded through the bank credit channel: the policy increased private-bank presence in underbanked districts; private banks then allocated credit disproportionately to ex-ante high-MRPK firms; this expansion reduced capital misallocation by compressing MRPK dispersion, though without immediate sales growth; and finally, ex-ante high-MRPK firms exhibited higher profitability and capital deepening but no significant rise in value added. The Indian experience thus illustrates both the promise and the limits of branch-expansion mandates—modest regulatory nudges can improve allocative efficiency in the short run, but durable productivity gains require reforms that realign the incentives of public-sector banks as well as private-sector banks. In particular, the absence of a comparable response from public-sector banks, despite their dominant role in the Indian banking system, meant that the reform's efficiency gains were narrow and temporary, highlighting the importance of bank ownership structure in shaping the effectiveness of financial sector reforms.

Layout We organize the paper as follows: Section 2 introduces the institutional background of the policy change, data sources, and measurement strategies. Section 3 outlines the empirical framework used to analyze the causal impact of the 2006 RBI policy and explores the underlying mechanism driving our causal findings. Section 4 presents robustness checks to validate our baseline results. Finally, Section 5 summarizes the key findings and proposes potential avenues for future research.

2 Institutional Background and Data

In this section, we discuss the institutional background of the Indian credit market and describe the methodology for our credit reallocation measures. Next, we present the data sources used to construct our credit reallocation and capital misallocation measures for the empirical analyses.

2.1 Bank Authorization Policy

Since the liberalization of the banking sector and economy in 1991, the Reserve Bank of India (RBI) has periodically distributed banking licenses to new banks and allowed private banks to operate retail banking in India. However, the banking sector remains dominated by public-sector banks. For example, around two-thirds of total deposits are held by PSBs, and PSBs conduct 70 percent of total loans and advances. Figure B.1 shows the share of different credits across time. The figure establishes that credit composition has not changed much over time for the banking sector. Figure B.1(c) shows an increase in the share of personal and services loans at the expense of industrial loans post-2006 among private-sector banks.

In 2006, the RBI introduced the Bank Authorization Policy (BAP), which standardized the bank branch application process by requiring banks wishing to expand or adjust bank branch operations to submit an application encompassing the strategic plans for the upcoming year. Pre-2006, bank branch proposals were approved case-by-case with no objective consideration of local credit markets. Simultaneously, the RBI classified whether local credit markets were being underserved by banks at the district level. Their measure was bank branch density, defined as the ratio of a district's population from the 2001 Census and the number of commercial bank branches in operation on March 31, 2005. A district was classified as underbanked if its bank branch density exceeded the national average. More precisely, the RBI defined an underbanked district as the following:

$$\frac{\#Bank\ Branches_{District}}{Population_{District}} < \frac{\#Bank\ Branches_{National}}{Population_{National}}.$$

We conjecture that BAP incentivized bank branch expansions to internalize the RBI's local market preferences, resulting in changes to local market credit dynamism characteristics that impact capital misallocation. However, we suggest these channels of effect are asymmetric by bank ownership type. As PSBs had a significant presence in underbanked districts pre-BAP, we expect PVBs to be more incentivized to expand into underbanked districts and respond more aggressively to preexisting bank competition by facilitating credit issuance to firms.

As the "underbanked" categorization was dependent on the national average, we have our treated (underbanked) and control (banked) districts as below and above the national average threshold. This allows us to use the national average branch density (scaled by population per million) – $\overline{Branch\ Per\ Capita}$ – as a threshold around which we have our treated and control district for our causal study. We look at those districts that are in the neighborhood of the national average (both above and below) to have comparable treatment and control districts. With $\overline{Branch\ Per\ Capita}$ acting as a threshold around which treated and control districts are chosen. We calculate the running variable as:

$$Runvar_d = Branch\ Per\ Capita_d - \overline{Branch\ Per\ Capita}$$

If we have $Runvar_d < 0$ then they are in the treated district and vice versa. Figure B.2 in the appendix shows the distribution of $Runvar$ with sufficient density around the threshold of the national average (scaled to 0). To ensure the pre-treatment comparability of treatment and control units, our primary sample is restricted to districts within a neighborhood of 20 (bank branches per million persons) around the national average threshold. The bandwidth is calculated using the optimal bandwidth estimation methodology of Calonico et al. (2020).[1] The optimal bandwidth of 20 around the threshold leaves us with a good sample of districts available for a study with 318 districts (216 underbanked and 102 banked districts). We also do robustness checks of our result with a bandwidth of 15 and 10 bank branches per capita and the complete sample.

Table A.1 captures the summary statistics for relevant variables for banked and under-

[1]We use the MSE optimal bandwidth around the threshold with a triangular kernel for the year 2005. Compared with other studies, the optimal bandwidth used by Kulkarni et al. (2023) is 15, and Young (2017) has a bandwidth of 3.5 thousand persons per branch to study the same policy intervention.

banked districts pre-BAP. The table shows that the mean difference in deposit and credit growth between banked and underbanked districts was insignificant. This suggests the treatment and control districts were similar before the policy change.

To further validate our design, in the robustness checks, we take steps to ensure that treated districts were genuinely affected by the reform. In particular, we focus only on underbanked districts where at least one new bank branch was actually opened after the policy took effect. Moreover, we require that these openings be linked to license applications submitted following the announcement of the policy, thereby excluding cases where branches were already in the pipeline beforehand. This refinement removes districts that might otherwise be misclassified as treated and strengthens our identification by ensuring that the estimated effects reflect the causal impact of policy-induced branch expansion rather than pre-existing trends or anticipatory behavior.

2.2 Data

Bank Branch Data We used proprietary data from the Reserve Bank of India's Basic Statistical Returns (BSR) to assess the impact of BAP on bank branch expansion. The BSR data set provides annual deposit and credit information for scheduled commercial banks and rural regional banks (RRBs) at the branch level that covers India. Bank identifiers further break down this data and provide sub-aggregated outstanding loans by sectoral classifiers for each bank branch. We used branch-level credit data restricted to industrial loans from 2001 to 2011. Further, this dataset reports new bank branches that have opened at the district level over time.

Firm Level Data The firm-level data in this study comes from the Prowess database maintained by the Center for Monitoring Indian Economy (CMIE). This database provides comprehensive information on a wide range of listed and unlisted companies. Prowess includes data on 1,500 items such as production, sales, profitability, liabilities, assets, and capital expenditures. It also offers detailed financial variables, including borrowing, bank loans, and secured and unsecured debt. The coverage of the Prowess database is extensive, accounting for 75 percent of corporate taxes and 95 percent of the excise duty collected by the Indian government (Kapoor et al., 2012). The database also categorizes firms by industry according to the 5-digit NIC classification. We use the length of data consistent with the BSR data. Since our main analysis compares firm-level responses in underbanked and banked districts, we use the reported district name variable in the Prowess database to identify whether firms are located in underbanked or banked districts.

In our causal analysis, we focus exclusively on manufacturing sector firms. Manufacturing firms are highly capital-intensive, making them particularly sensitive to inefficiencies in resource distribution compared to other sectors. Moreover, the manufacturing sector offers a more standardized framework for analyzing MRPK dispersion, as variations in production processes and capital utilization are relatively comparable within this sector. This focus allows us to leverage the detailed and comprehensive nature of the Prowess database to assess capital misallocation and its underlying mechanisms with greater precision.

Firm-Bank Loan Matched Data The Ministry of Corporate Affairs (MCA) of the Indian Central Government tracks all collateralized loans issued to registered firms by both bank and non-bank financial institutions. This dataset includes the borrower's identity, loan value, issuance date, creditor's identity, and the firm's unique corporate identification number (CIN). Firms must report loan issuance within 20 days using a unique charge number, notify the MCA upon full repayment, and report any modifications. Covering 40-50 percent of outstanding commercial bank loans in India from 2001-2016 (see Chakraborty and Ritadhi, 2022), this data, with its firm identifiers, is well-suited for merging with the Prowess database and analyzing collateralized loans by bank ownership type.

For the firm-banked matched data from 2001 to 2011, the dataset includes 144,885 new loans issued by public-sector, private-sector, foreign banks, and non-banking financial institutions (NBFIs). We focus on public and private-sector bank loans, resulting in 97,171 loan issuances. Since we restrict our firm-level sample to manufacturing sector firms, we select a sub-sample of 34,936 loan issuances to manufacturing firms matched with the CMIE Prowess dataset, covering approximately 36 percent of the loans issued by public and private-sector banks. We use the MCA bank-level secured loan data as a proxy for total loans, as secured loans are nearly identical to total commercial loans (over 95 percent in the CMIE Prowess dataset).

Table A.2 documents the summary statistics of loan amounts by creditor type for all firm-bank loan matched data (Panel A) and the merged sample of firm-bank loan matched data (Panel B) for the 2001-2011 period, respectively. We have several key observations. First, in both panels, the mean loan amounts for public-sector banks are higher than those for private-sector banks. Second, the merged sample shows higher median loan amounts for both types of banks. Table A.3 documents the summary statistics on how firm borrower characteristics differ by bank ownership (Public-sector vs. Private-sector banks).

We observe that Private-sector banks tend to lend to firms with higher MRPK, assets, sales, and cash profits than Public-sector banks, and we will analyze this systematically through the lens of policy in the empirical section.

2.3 Measurement

Measurement of Capital Misallocation We use the firm-level marginal revenue product of capital (MRPK), a sufficient statistic for capital misallocation due to its ability to directly represent how efficiently capital is allocated across firms. Theoretically, MRPK should equalize across firms in an economy with optimal capital allocation. Any observed dispersion in MRPK signals deviations from this optimal state, indicating misallocation where some firms hold excess capital while others are undercapitalized relative to their productivity.

In our analysis, we measure firm-level MRPK using the Prowess database, following the standard approach in the literature, where MRPK is proxied by firm-level total sales per total capital (see Hsieh and Klenow, 2009; Sraer and Thesmar, 2023b; Bau and Matray, 2023). This proxy captures firm-specific variations in capital productivity, providing a robust measure of allocative inefficiency. This approach aligns with the theoretical framework of misallocation studies, where MRPK dispersion is a widely recognized metric for inefficiencies in capital distribution.

Table A.4 documents the summary statistics of the logarithm of firm-level MRPK in our sample across underbanked and banked districts before and after the bank branch expansion policy. We also provide several stylized facts on the trends in capital misallocation measures in the Indian economy. First, Figure B.3 presents the histogram and kernel density of the firm-level MRPK measure. Both figures show that firm-level MRPK exhibits substantial variation across districts and time, which provides good heterogeneity for our causal analysis.

Moreover, Figure B.4(a) displays the annual mean logarithm of firm-level MRPK (total sales per total capital) for underbanked and banked districts over time. A stable pattern is observed for both district types before the policy, whereas an increasing trend emerges after the policy's implementation. This suggests potentially significant heterogeneous effects of the policy on firm-level MRPK between underbanked and banked districts. Next, we analyze how the dispersion in firm-level MRPK behaves over time. We calculate capital misallocation based on the dispersion of MRPK across firms within districts and years. Figure B.4(b) shows a parallel trend for both district types before the policy, with capital misallocation exacerbating over time for both underbanked and banked districts. However, we observe a decline in capital misallocation during the policy fiscal year (2006) in underbanked districts, while the trend remains similar in banked districts. This observation suggests a potential impact of the policy on capital misallocation, which we investigate more systematically and causally in the empirical section.

3 Empirical Evidence

In this section, we investigate the mechanisms underlying the observed causal effects of the bank authorization policy, focusing on the roles of bank ownership and firm heterogeneity. Specifically, we examine how public and private sector banks differed in their responses to the policy, particularly in their lending behaviors and credit allocation to firms with varying MRPK levels. By analyzing trends in branch expansion, lending practices, and firm-level outcomes, we identify the key drivers of changes in capital allocation and efficiency in underbanked districts.

3.1 Policy Impact on Bank Branches by Bank Ownership

We look at the impact of the policy on the number of branch openings post-policy. Figure B.5 captures the evolution of the total number of bank branches within our sample over time for all banks and banks with different ownership types. In the figure, we compare the evolution between underbanked and banked districts that lie in the bandwidth of *20 bank branches per million* around the national average threshold. We find in Figure B.5 the total number of bank branches in underbanked districts disproportionately increased while the trend remained the same for banked districts post-policy. When we look at public-sector banks, Figure B.5 highlights no visible change in trend pre and post-policy, and we see similar differences over time. In the case of private-sector banks, we find a significant effect of the policy seen in B.5 with the number of bank branches increasing significantly in underbanked districts compared to banked districts around the national average threshold cutoff, suggesting the expansion of bank branches in underbanked districts can be attributed to PVBs.

To add rigor to our argument, we use a difference-in-differences methodology using the following equation:

$$\begin{aligned} Y_{dt} = \alpha + \beta_1 \text{Post}_t + \beta_2 \text{Underbanked}_d + \beta_3 (\text{Underbanked}_d \times \text{Post}_t) \\ + \Gamma X_{dt} + \varepsilon_t + \phi_d + \upsilon_{st} + \epsilon_{dt} \end{aligned} \tag{1}$$

where Y_{dt} represents the log of number of bank branches by total banks or by different ownership for district d and year t. Post$_t$ denotes a dummy variable indicating whether time t is post-treatment (i.e., $t \geq 2006$), taking values of 1 for post-treatment periods and 0 for pre-treatment periods. Underbanked$_d$ is a dummy variable indicating whether district d is an underbanked district, which takes a value of 1 for underbanked districts and 0 for banked districts. The coefficient of our interest is β_3, which shows the effect of the policy change on under-banked districts vis à vis banked district with respect to the number of branches (in percentage term). X_{dt} represents district-level control variables, such as the logarithm of the population and the logarithm of the number of bank branches (for some previous period), which are interacted with a linear-time trend. Furthermore, we include year (ε_t), district (ϕ_d), and state-year fixed effects (υ_{st}) to control for unobserved related factors. Standard errors are clustered at the district level.

Table 1: Difference-in-Differences Regression Result for Bank Branch Expansion by Bank Ownership

	Total Branches			PSB Branches			PVB Branches		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Underbanked × Post	-0.001	-0.004	0.005	-0.010	-0.015	-0.008	0.101*	0.131***	0.163***
	(0.014)	(0.015)	(0.016)	(0.012)	(0.013)	(0.014)	(0.056)	(0.044)	(0.047)
Distirct FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State×Year FE	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
District Trends	No	No	Yes	No	No	Yes	No	No	Yes
R^2	0.993	0.995	0.995	0.994	0.996	0.996	0.932	0.964	0.965
Observations	3,494	3,471	3,377	3,494	3,471	3,377	2,453	2,411	2,365

Note: The table above is a Difference-in-Differences regression for the log of number of bank branches as specified in equation (1) to assess the effect of policy change on refined underbanked districts vis à vis banked districts. Here, we are only looking at the average treatment effect of the treated on overall banks and, then, separately by bank ownership that lies in the bandwidth of *20 bank branches per million* around the national average threshold. The dependent variables are the log of total number of branches at the district level. It is regressed against $Post_t \times Underbanked_d$ and $Underbanked_d$ takes the value of 1 if the district is underbanked, whereas $Post_t$ takes a value of 1 if the year is on and after 2006. The first columns for each variable refer to the results with limited fixed effects, while the last column constitutes all the fixed effects and other controls. The study spans data from fiscal year 2001 to 2011. The analysis is at the district level, is annual, and covers all Indian districts as per Census 2001 in an unbalanced panel. Standard errors (in parentheses) are clustered at the district level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 1 shows the difference in the differences results for all banks as well as by their type of ownership. We do not find any significant increase in branches across all branches in underbanked districts compared to banked districts post-policy. Additionally, we find no significant impact with respect to public-sector banks, as shown in column (6). We find a positive and significant effect on private sector banks', showing a 16.3% increase in bank branches post policy for underbanked districts vis à vis banked districts as shown in column (9) of Table 1. Our differences-in-differences results corroborate the trend that we have found in Figure B.5 where branch increase in "underbanked" districts is driven primarily by private sector banks.

We further use an event study plot to document the trajectory of the treated group (i.e., "underbanked districts") over time for number of bank branches. The event study that we have used is a standard event study plot, and the equation for the event study regression is the following:

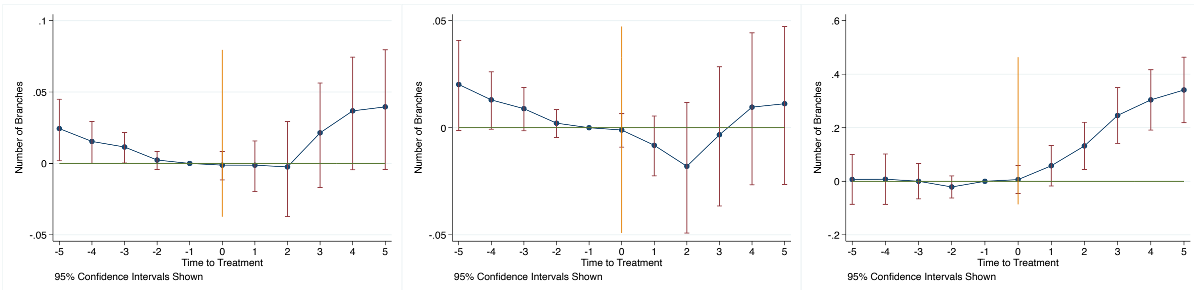
$$Y_{dt} = \sum_{l=-k}^{m} \beta_l \mathbb{1}\{Underbanked_d\} \times \mathbb{1}\{t = 2006 + l\} + \Gamma X_{dt} + \varepsilon_t + \phi_d + \upsilon_{st} + \epsilon_{dt} \tag{2}$$

Here $k > 0$ and $m > 0$, we use the standard fixed effects as used earlier. Using an event study plot, we try to see how β_l evolves over time with a particular focus on the post-policy change. To be more precise, we want to reject the hypothesis that there are pre-trends for number of bank branches between the two types of districts, and secondly, we want to see the evolution of β_l post-policy change to document whether the effect was transitory or permanent.

Using the standard difference-in-differences regression, we have seen an average effect of bank branch expansion policy on underbanked districts vis à vis banked districts for log of total number of branches, and we saw a consistent result at the district level. We further want to consolidate our result by using an event study plot at the district level and observing the trajectory of the estimate over time.

Figure 1 shows the evolution of β_l over time for the log of the number of branches in all banks and banks by their type of ownership. We see a clear positive and significant effect after the policy intervention on the number of bank branches for private sector banks, as shown in the panel (c) of Figure 1. This is in line with what we get in Figure B.5 and Table 1, i.e., bank branch expansion is primarily driven by private sector banks in underbanked districts. On the other hand, we do not find any significance in the post-policy coefficient β_l for all banks and public sector banks, as shown in Figure 1 panel (a) and (b), respectively. Hence, to sum up, we can conclude using trend plots, difference in differences results, and event studies that private sector banks are the ones that are leading to the increase in bank branch expansion in underbanked districts, with no such effect for public sector banks.

Figure 1: Dynamic DiD - Number of Bank Branches by Bank Ownership



(a) All Banks (b) Public Banks (c) Private Banks

Note: The figure shows the dynamic evolution of the log of number of branches by different types of banks over time for our refined underbanked and banked districts that lie in the bandwidth of *20 bank branches per million* around the national average threshold. The regression model includes district-level control variables (the logarithm of the population and the logarithm of the number of bank branches interacted with a linear time trend), year, district, and state-year fixed effects. Standard errors are clustered at the district level. The sample period spans from 2001 to 2011, with 2006 as the fiscal year of $t = 0$, corresponding to the implementation of the bank branch expansion policy. Confidence intervals are calculated at the 95% level.

3.2 Policy Impact on Capital Misallocation and Firm Outcomes

This part investigates the causal impact of the bank branch policy change on capital misallocation. Given the plausibly exogenous policy change, we employ a difference-in-differences (DiD) framework to estimate the causal effect using the following regression specification:

$$\begin{aligned} Y_{idt} = \alpha &+ \beta_1 \text{Post}_t + \beta_2 \text{Underbanked}_d + \beta_3 (\text{Underbanked}_d \times \text{Post}_t) \\ &+ \beta_4 (\text{Underbanked}_d \times \text{Post}_t \times I^{HighMRPK}) \\ &+ \Gamma X_{dt} + \varepsilon_t + \phi_d + \upsilon_{st} + \epsilon_{idt} \end{aligned} \tag{3}$$

where Y_{idt} represents the logarithm of firm i's MRPK (total sales per total capital) in district d and year t. Post_t denotes a dummy variable indicating whether time t is post-treatment (i.e., after the year 2006), which takes a value of 1 for post-treatment periods and 0 for pre-treatment periods. Underbanked_d is a dummy variable indicating whether district d is an underbanked district, which takes a value of 1 for underbanked districts and 0 for banked districts. Since we are interested in how the policy affects capital misallocation between underbanked and banked districts, we examine how firms with different ex-ante MRPK levels performed before and after its implementation. For instance, if firms with higher ex-ante MRPK experience a decrease (increase) in their MRPK after the policy relative to firms with lower ex-ante MRPK, this would suggest a decline (increase) in capital misallocation. To investigate this, we construct a dummy variable, I^{HighMRPK}, which equals one if a firm's ex-ante MRPK is higher than the industry mean MRPK during the pre-policy period and zero otherwise. We interact I^{HighMRPK} with Post_t and Underbanked_d to analyze its impact in conjunction with the policy change. X_{dt} represents district-level control variables, such as the logarithm of the population and the logarithm of the number of bank branches interacted with a linear time trend. Additionally, we include year (ε_t), district (ϕ_d), and state-year fixed effects (υ_{st}) to control for unobserved related factors.

Identification Assumptions The identification of the causal effects in our empirical framework relies on several key assumptions. First, the parallel trends assumption requires that the trajectory of firm-level MRPK (Y_{idt}) would have been similar for firms in underbanked and banked districts absent of the policy. Second, there should be no differential pre-trends based on I^{HighMRPK}, meaning firms with higher ex-ante MRPK ($I^{\text{HighMRPK}} = 1$) in underbanked and banked districts must not exhibit different trends in MRPK relative to lower MRPK firms before the policy, conditional on the controls. Third, the correct specification of controls requires that district-level variables (X_{dt}) and fixed effects ($\varepsilon_t, \phi_d, \upsilon_{st}$) adequately capture observable and unobservable factors that could influence MRPK trends across districts and over time. Fourth, the exogeneity of policy implementation assumes that the policy and the classification of districts as underbanked or banked are not systematically related to unobserved factors that directly affect MRPK trends. Fifth, there must be no spillover effects, ensuring that the policy's impact on MRPK in underbanked districts does not influence outcomes in banked districts or vice versa. Finally, the exogeneity of I^{HighMRPK} requires that the classification of firms based on their ex-ante MRPK is not correlated with unobserved factors driving MRPK trends. These assumptions collectively ensure that the estimated coefficients can be interpreted as the causal effects of the policy.

A potential concern with our baseline identification strategy is the possibility of firms operating multiple establishments across different districts. In such cases, credit shocks or policy-induced changes in one district may spill over into another, thereby undermining the strict geographic exogeneity required for causal inference. To address this, we also conduct a refined analysis in the robustness check by restricting the sample to firms plausibly operating within a single district. This single-district approach strengthens our identification by ensuring that firm outcomes are directly tied to the local credit market where the policy intervention occurred, thereby mitigating concerns of cross-district spillovers.

Table 2 documents the difference-in-differences estimates. First, we find some evidence of a positive impact on MRPK for underbanked districts after the policy, but this effect is not robust across all specifications. However, we document that the policy alleviates capital misallocation in underbanked districts, as it negatively and significantly impacts the MRPK of high ex-ante MRPK firms after the policy. We find that for firms with initially high marginal revenue products of capital (MRPK), the policy reduces MRPK by around 60 percent (Column 8 of Table 2) relative to firms with low MRPK. This suggests that the dispersion in MRPK across firms in underbanked districts diminishes, resulting in lower capital misallocation.

Table 2: Difference-in-Differences Regression Result for Capital Misallocation

	(1) Log MRPK	(2) Log MRPK	(3) Log MRPK	(4) Log MRPK	(5) Log MRPK	(6) Log MRPK	(7) Log MRPK	(8) Log MRPK
Underbanked × Post	0.0403	0.0405	0.00131	-0.0836	-0.0178	-0.0188	-0.0296	-0.0828
	(0.0923)	(0.0916)	(0.0838)	(0.0975)	(0.0689)	(0.0671)	(0.0667)	(0.0777)
Underbanked × Post × $I^{HighMRPK}$					-0.750***	-0.744***	-0.631***	-0.591**
					(0.208)	(0.208)	(0.178)	(0.187)
District FE	No	No	Yes	Yes	No	No	Yes	Yes
Year FE	No	Yes	Yes	Yes	No	Yes	Yes	Yes
State×Year FE	No	No	No	Yes	No	No	No	Yes
District Trends	No	No	No	Yes	No	No	No	Yes
R^2	0.00965	0.0152	0.109	0.111	0.224	0.230	0.303	0.304
Observations	15,759	15,759	15,754	15,067	12,890	12,890	12,890	12,415

Note: This table presents the difference-in-differences estimates of the regression model outlined in equation (3). Post_t indicates whether time t is post-treatment (i.e., after the year 2006), with a value of 1 for post-treatment periods and 0 otherwise. Underbanked_d is a dummy variable indicating whether district d is underbanked (1 for underbanked districts, 0 otherwise). I^{HighMRPK} is a dummy variable set to 1 if a firm's ex-ante MRPK before the policy exceeds the industry mean MRPK within the year. District-level control variables are the logarithm of the population and the logarithm of the number of bank branches, which interact with a linear time trend. Here, we are looking at the average treatment effect on the treated districts that lie within the bandwidth of *20 bank branches per million* around the national average threshold and compared with the control districts. Standard errors (in parentheses) are clustered at the district level. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

As a next step, to explore the heterogeneous dynamic responses of firms with different MRPK, we use the dynamic DiD and event study approach, employing the following specification:

$$Y_{idt} = \sum_{k=-5}^{+5} \beta_k (\mathbb{1}\{\text{k Years to Policy}\} \times \text{Underbanked}_{d}) + \Gamma X_{dt} + \varepsilon_t + \phi_d + \upsilon_{st} + \epsilon_{idt} \qquad (4)$$

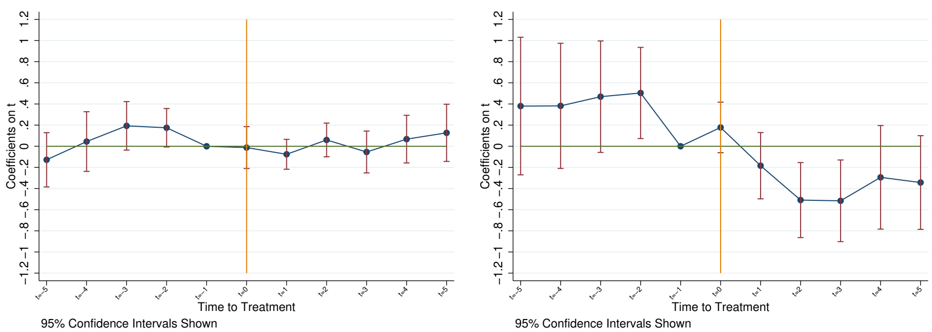
where $\mathbb{1}\{\text{k Years to Policy}\}$ is an indicator function that takes the value of 1 if the difference between a particular year and the policy fiscal year of 2006 is k, with $k \in [-5, 5]$. Other variables are the same as in the regression model (3). To investigate the channel of capital misallocation, we split the firms based on those with lower and higher ex-ante MRPK before the policy and ran the regression specification (4) within each group.

Figure 2 documents several important results. First, we confirm that there is no evidence of pre-trends, as the outcome variable—logarithm of firm-level MRPK—is indistinguishable between exposed (underbanked) and control (banked) districts for both ex-ante low MRPK firms and ex-ante high MRPK firms.[2] Second, we show that the policy does

[2]The coefficient at $t = -2$ for ex-ante high MRPK firms in Panel (b) is statistically significant. However,

not lead to any differential impacts on MRPK for firms with lower ex-ante MRPK between underbanked and banked districts. In contrast, it causes firms with higher ex-ante MRPK to experience a drop in their MRPK in underbanked districts during the few years after the policy. However, this effect is short-lived, lasting only 2-3 years. Therefore, our evidence reveals that firms with higher ex-ante MRPK before the policy experienced a decline in their MRPK after the policy in underbanked districts. This results in lower dispersion of MRPK across firms and, hence, a temporary reduction in capital misallocation.

Figure 2: Dynamic DiD - Capital Misallocation



(a) Ex ante Low MRPK Firms (b) Ex ante High MRPK Firms

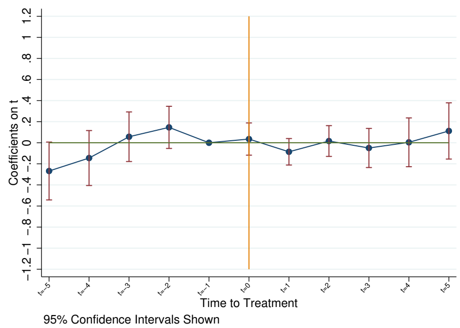
Note: This figure presents the dynamic difference-in-differences estimates of the regression model outlined in equation (4). We categorize firms as ex-ante high (low) MRPK if their average MRPK is above (below) the 2-digit industry-level mean from 2001-2005. The regression model includes district-level control variables (the logarithm of the population and the logarithm of the number of bank branches interacted with a linear time trend), year, district, and state-year fixed effects. Here, we are looking at the average treatment effect on the treated districts that lie within the bandwidth of *20 bank branches per million* around the national average threshold and compared with the control districts. Standard errors are clustered at the district level. The sample period spans from 2001 to 2011, with 2006 as the fiscal year of $t = 0$, corresponding to the implementation of the bank branch expansion policy. Confidence intervals are calculated at the 95% level.

3.2.1 Policy Impact on the Components of Capital Misallocation

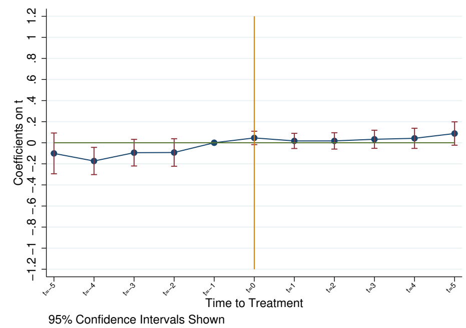
To examine how ex-ante low- and high-MRPK firms respond to the policy that reduces capital misallocation in underbanked districts, we analyze the heterogeneous responses in sales and capital among these two types of firms, as these variables form the basis of our capital misallocation measure.

a joint F-test for the pre-treatment coefficients fails to reject the null hypothesis of no differential pre-trends, with an F-statistic of 1.458 and a p-value of 0.223. Thus, we confirm that the parallel trend assumption is satisfied based on the statistical test.

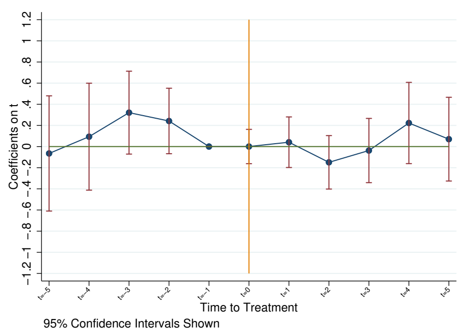
Figure 3: Dynamic DiD - Firm Capital and Sales



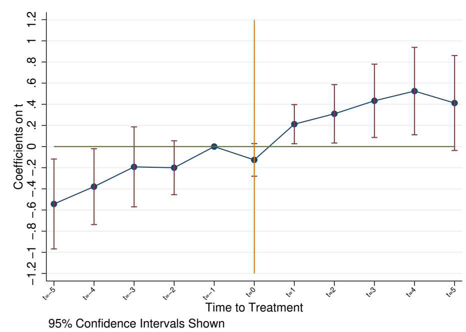
(a) Ex ante Low MRPK - Sales



(b) Ex ante Low MRPK - Capital



(c) Ex ante High MRPK - Sales



(d) Ex ante High MRPK - Capital

Note: This figure presents the dynamic difference-in-differences estimates on firm capital and sales. We categorize firms as ex-ante high (low) MRPK if their average MRPK is above (below) the 2-digit industry-level mean from 2001-2005. The regression model includes district-level control variables (the logarithm of the population and the logarithm of the number of bank branches interacted with a linear time trend), year, district, and state-year fixed effects. We estimate the average treatment effect on the treated districts within a bandwidth of *20 bank branches per million* around the national average threshold, comparing them with control districts. Standard errors are clustered at the district level. The sample period spans from 2001 to 2011, with 2006 as the fiscal year of $t = 0$, corresponding to the bank branch expansion policy implementation. Confidence intervals are calculated at the 95% level.

Figure 3 presents a dynamic DiD analysis examining the impact of the bank branch policy on sales and capital patterns among manufacturing firms, categorized by their MRPK levels before the policy. We find that ex-ante low MRPK firms display relatively stable sales patterns, with only modest fluctuations around the policy implementation period. Post-policy, their capital investments show a slight but statistically insignificant upward trend, suggesting limited responsiveness to the policy. In contrast, ex-ante high MRPK firms in underbanked districts exhibit more dynamic responses. While no significant increase in sales is observed post-policy, their capital accumulation shows a pronounced and statistically significant upward trend. This indicates that high MRPK firms could better capitalize on the opportunities provided by the policy, substantially enhancing their capital base. Hence, the increase in capital without a corresponding significant rise in sales explains why the MRPK of ex-ante high MRPK firms declined post-policy. The added capital base dilutes their marginal productivity of capital, leading to a reduction in MRPK.

Another dimension to consider is that the policy may affect firm outcomes beyond capital misallocation. To provide a broader perspective, in the robustness checks of Section 4, we extend the analysis to additional key firm-level outcomes, examining the policy's effects on capital deepening, value added, and profit margins. We find that ex-ante high-MRPK firms in underbanked districts experience a significant increase in capital per employee and higher profit margins following the policy, but no corresponding change in gross value added. By contrast, ex-ante low-MRPK firms do not exhibit meaningful changes in any of these outcomes. Taken together, these results suggest that, in addition to reducing capital misallocation, the policy fostered capital deepening and improved profitability among ex-ante high-MRPK firms, though without a parallel rise in value added.

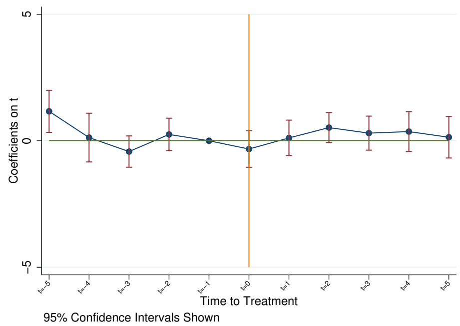
Building on our benchmark results—declining capital misallocation, increases in capital per employee and profit margins, and no corresponding rise in sales or value added—we next explore heterogeneity by examining the role of bank ownership. To shed light on the underlying mechanism, we use matched bank–firm data to analyze how ownership structures mediate the policy's impact on capital allocation. We further investigate how different types of firms benefiting from bank lending respond to the policy, thereby driving the observed changes in misallocation.

3.3 Policy Impact on Credit Allocation by Bank Ownership

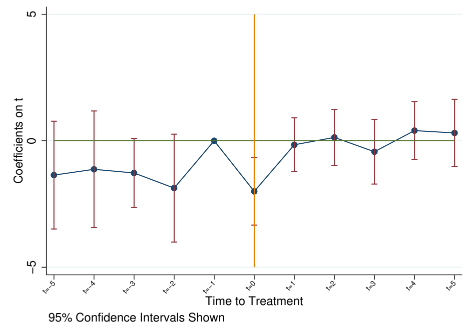
To examine the role of bank ownership in our Difference-in-Differences (DiD) analysis, we analyze the relationship between bank ownership, capital misallocation, and borrower firm characteristics across underbanked and banked districts.

Utilizing the matched bank-firm data, Figure 4 provides a detailed comparison of how public sector banks (PSBs) and private sector banks (PVBs) adjusted their lending patterns in response to a bank branch policy, focusing specifically on manufacturing firms categorized by their ex-ante MRPK before policy implementation. Notably, Figure 4(f) presents significant evidence that PVB lending to high-MRPK firms increased markedly following the policy. Specifically, ex-ante high-MRPK firms in underbanked districts received significantly larger loans from private sector banks post-policy than the control group.

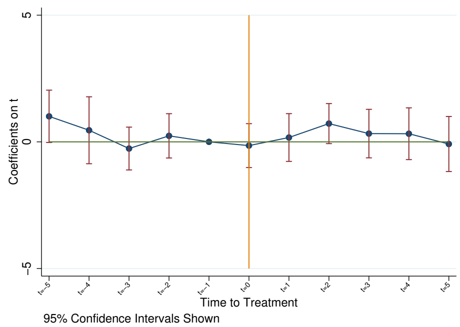
Figure 4: Dynamic DiD - Public and Private Sector Bank Loans



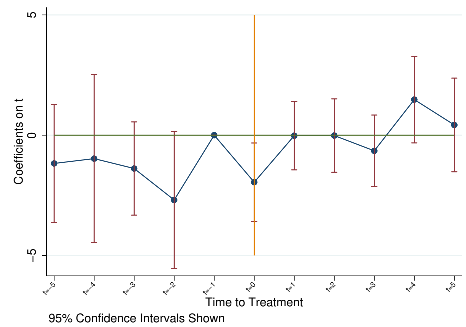
(a) All - PSB Loans



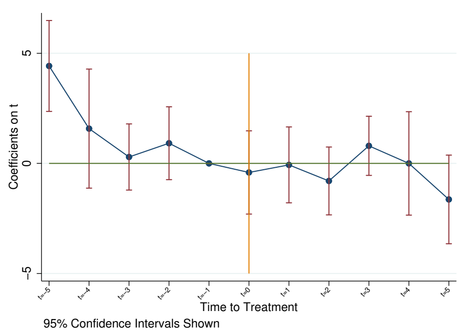
(b) All - PVB Loans



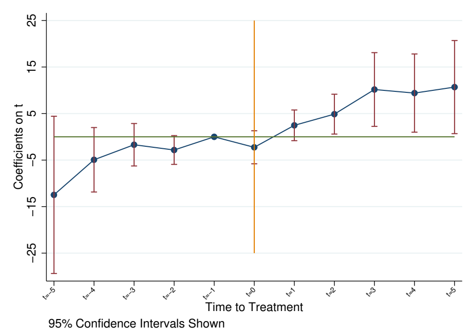
(c) Ex ante Low MRPK - PSB Loans



(d) Ex ante Low MRPK - PVB Loans



(e) Ex ante High MRPK - PSB Loans



(f) Ex ante High MRPK - PVB Loans

Note: We categorize firms as ex-ante high (low) MRPK if their average MRPK is above (below) the 2-digit industry-level mean from 2001-2005. The regression model includes district-level control variables (the logarithm of the population and the logarithm of the number of bank branches which are interacted with a linear time trend), year, district, and state-year fixed effects. Loan amounts are deflated. Here, we are looking at the average treatment effect on the treated districts within the bandwidth of *20 bank branches per million* around the national average threshold and compared with the control districts. Standard errors are clustered at the district level. The sample period spans from 2001 to 2011, with 2006 as the fiscal year of $t = 0$, corresponding to the implementation of the bank branch expansion policy. Confidence intervals are calculated at the 95% level.

In contrast, ex-ante low-MRPK firms in underbanked districts did not experience significant changes in loans from public or private sector banks after the policy. This indicates that PVBs underwent a temporary shift in their lending practices, prioritizing higher MRPK firms, whereas a similar shift is not observed for PSBs. This finding is particularly significant as it builds on our earlier evidence that the policy reduced capital misallocation in underbanked districts. This reduction was driven by a decline in the MRPK of ex-ante (pre-policy) high MRPK firms, accompanied by increased credit growth facilitated primarily by PVBs.

By integrating these findings, we suggest that the temporary adjustment in PVB lending practices—providing more credit to relatively higher MRPK firms in underbanked districts—is a key mechanism behind the observed decline in MRPK for these firms post-policy. This targeted allocation of credit growth disproportionately benefited high MRPK firms, which, in turn, contributed to the observed reduction in capital misallocation. Additionally, this pattern highlights how the increased credit growth in underbanked districts was not evenly distributed but instead channeled predominantly toward higher MRPK firms. This reallocation of resources played a crucial role in improving capital efficiency in these areas.

To understand the characteristics of ex-ante high- and low-MRPK firms, we perform a probit regression to identify key predictors of whether a manufacturing firm is classified as ex-ante high MRPK (i.e., having an MRPK above the industry mean) during the pre-policy period (2001–2005). Table 3 shows that log assets (firm size) have a consistently positive and highly significant association with being a high MRPK firm, indicating that larger firms are more likely to exhibit higher capital productivity. Similarly, firms with higher cash ratios are positively associated with high MRPK status, suggesting that liquidity supports efficient capital use. Conversely, the leverage ratio has a negative and significant relationship with the high MRPK classification, implying that heavily indebted firms tend to have lower capital productivity. Firm age is another strong predictor, with older firms more likely to be classified as high MRPK, reflecting the potential benefits of accumulated experience and stability in resource allocation. These relationships are robust across specifications and control for district, state, and industry fixed effects. The findings highlight the role of firm size, liquidity, and financial structure in determining capital productivity, offering critical insights into the characteristics of firms most likely to benefit from policies targeting capital misallocation.

Table 3: Probit Regression - Predictors of Ex ante High and Low MRPK Firms

	(1) $I^{HighMRPK}$	(2) $I^{HighMRPK}$	(3) $I^{HighMRPK}$	(4) $I^{HighMRPK}$	(5) $I^{HighMRPK}$
Log Assets	0.116***				0.150***
	(0.0238)				(0.0313)
Cash Ratio		0.0137			0.102**
		(0.0297)			(0.0366)
Leverage Ratio			-0.171***		-0.205***
			(0.0339)		(0.0417)
Age				0.192***	0.221***
				(0.0570)	(0.0650)
District FE	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes
Observations	1,626	1,402	1,561	1,622	1,359

Note: The dependent variable I^{HighMRPK} is a dummy variable set to 1 if a firm's ex-ante MRPK before the policy exceeds the 2-digit industry mean MRPK within the year during the 2001-2005 (pre-policy) period. The explanatory (predictor) variables are the firm-level average of log assets, cash ratio, leverage ratio, and age during the 2001-2005 (pre-policy) period. The cash ratio is measured as firm-level cash balances per total assets, and the leverage ratio is measured as firm-level total borrowings per total assets. Age is measured as the difference between a corresponding year and a firm's incorporation year. All related variables are deflated using the GDP deflator. Here, we keep the districts that lie within the bandwidth of *20 bank branches per million* around the national average threshold. Standard errors (in parentheses) are clustered at the firm level. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

When interpreting our DiD results—showing that ex-ante high MRPK firms in underbanked districts receive disproportionately more loans from private sector banks—through the lens of the results in Table 3, we argue that private sector banks are more likely to lend to ex-ante high MRPK firms that are larger, cash-rich, less leveraged, and more mature. This preference for cash-rich and lower-leveraged firms suggests that private-sector banks prioritize lower-risk borrowers. Similarly, the focus on larger and more mature firms reflects a tendency to favor more established and financially stable businesses. This lending behavior emphasizes private sector banks' cautious and selective approach in targeting firms with strong financial fundamentals. In our context, private-sector banks contribute to reducing capital misallocation by prioritizing loans to high MRPK firms with stable financial profiles. These firms are better positioned to use the additional capital productively, enhancing overall economic efficiency in underbanked districts.

Combined with findings from the previous section, these results highlight that ex-ante high MRPK firms in underbanked districts, which received higher private sector bank loans post-policy, significantly increased their capital but not their sales. This shift in capital allocation contributes to a decline in their MRPK and, consequently, a reduction in capital misallocation in underbanked districts. This evidence highlights the pivotal role of private sector banks in driving improvements in capital efficiency through targeted credit allocation.

4 Robustness Checks and Further Analysis

In this section, we perform robustness checks and further analysis for our main findings regarding the policy's impacts and its underlying mechanisms. Specifically, we examine a refined identification strategy, additional firm outcomes, and the policy effects under different bandwidth selections, including no bandwidth (i.e., all districts), a bandwidth of 15, and a bandwidth of 10.

Refined Identification Strategy A potential concern with our baseline identification strategy is the possibility of firms operating multiple establishments across different districts. In such cases, credit shocks or policy-induced changes in one district may spill over into another, thereby undermining the strict geographic exogeneity required for causal inference. To address this, we refine our identification by restricting the analysis to firms plausibly operating within a single district.

Specifically, we utilize our matched firm-bank loan data and focus on firms that borrow exclusively from bank branches located within their own district throughout the sample period. While the Prowess dataset does not directly indicate whether a firm operates in multiple districts, we argue that single-district borrowing restriction serves as a reasonable proxy for single-district firms. Firms that borrow only from local bank branches are more likely to be locally embedded and thus directly exposed to district-level credit supply conditions.

Additionally, we ensure that our treated districts were genuinely exposed to the policy. We accomplish this goal in two steps. First, we restrict the sample to underbanked districts that experienced at least one bank branch opening following the policy. Second, we make sure that these new branches are opened with branch license applications made after the policy announcement. Effectively, we exclude underbanked districts with new branch license applications made before the policy. This refinement eliminates placebo-treated districts and strengthens the interpretation of our estimates as capturing the causal impact of actual policy exposure.

Figure B.6 presents the benchmark event study results for capital misallocation under the refined identification strategy, restricting the analysis to single-district firms and underbanked districts that experienced realized branch expansion. We again find that ex-ante high MRPK firms exhibit a significant decline in their MRPK following the policy in underbanked districts, while ex-ante low MRPK firms show no significant change. This pattern implies a reduction in the dispersion of MRPK, i.e., a decline in capital misallocation in these districts. Overall, these results confirm that our main findings on capital misallocation remain robust to alternative identification assumptions.

Additional Firm Outcomes Another dimension to consider is that the policy may also affect firm outcomes beyond capital misallocation. To provide a broader perspective, we extend our analysis to key indicators of firm dynamics, examining the effects of the policy on capital deepening, value added, and profit margins. Consistent with the previous robustness check, we apply the refined identification strategy, focusing on single-district firms and districts that experienced realized branch expansion.

The results, shown in Figure B.7, indicate that ex-ante high MRPK firms experience a significant increase in capital per employee and higher profit margins in underbanked districts following the policy. However, there is no significant change in their gross value added. In contrast, ex-ante low MRPK firms do not exhibit meaningful changes in any of these outcomes. These findings suggest that, in addition to reducing capital misallocation, the policy also induces capital deepening and improved profitability among ex-ante high MRPK firms, though without a corresponding rise in value added.

Trends and DiD for Bank Branches by Bank Ownership Figure B.8 illustrates the trends in the number of branches by bank ownership across various bandwidths. In line with our benchmark results, the data reveal a similar pattern: the number of branches of private sector banks in underbanked districts increases substantially. At the same time, no significant change is observed in the trend for public sector banks. Figure B.9 presents the DiD estimates for the number of branches by bank ownership, further confirming that private sector bank branches experience a significant increase in underbanked districts. In contrast, no such evidence is found for public sector banks, even after controlling for various district-level control variables and applying multiple fixed effects. We also find that when we look at the complete sample with no bandwidth around the threshold, the result is still consistent, as can be seen in Table A.5, showing significant effect on private bank branches and no such effects on public sector banks.

DiD for Capital Misallocation Tables A.6, A.7, and A.8 present the benchmark DiD estimates, while Figure B.10 shows the dynamic DiD estimates from our benchmark analysis of the policy's impact on capital misallocation, considering three bandwidth selections: no bandwidth, a bandwidth of 15, and a bandwidth of 10. The tables confirm the policy's impact on capital misallocation across different bandwidth selections, and the figure demonstrates that the parallel trends assumption holds consistently across all bandwidth selections. Additionally, we find that ex-ante high-MRPK firms in underbanked districts experience a downward trend in their MRPK post-policy, whereas ex-ante low-MRPK firms show no significant change in their MRPK.

DiD for Heterogeneous Impacts by Bank Ownership Figure B.11 presents the DiD estimates for private and public sector bank loans to ex-ante low and high MRPK firms across the full sample, with no bandwidth applied.[3] Our analysis again shows that ex-ante high MRPK firms in underbanked districts receive significantly higher private-sector lending. However, we find no significant change in the loan allocation for ex-ante low MRPK firms, regardless of whether the loan comes from private or public sector banks.

DiD for Heterogeneous Firm Responses Figures B.12 and B.13 display the DiD estimates for capital and sales responses, respectively, for ex-ante low and high MRPK firms across different bandwidths. Consistent with our benchmark results, ex-ante high MRPK firms in underbanked districts experience a significant increase in capital post-policy while their sales remain relatively stable. In contrast, we again find no significant changes in capital or sales for ex-ante low MRPK firms across all bandwidth selections.

Overall, through a comprehensive set of robustness checks, including variations in bandwidth selections, we demonstrate that our benchmark findings regarding the policy's impact on capital misallocation and its underlying mechanisms remain robust across different specifications. These additional checks enhance the reliability of our results and provide further assurance that the observed effects are not sensitive to the choice of bandwidth.

5 Conclusion

Our study investigates the efficacy of banking reforms within a developing economy context. We focus on how lending behaviors and financial intermediaries' structural incentives shape the credit flow toward firms and how this process affects capital allocation. Using the Reserve Bank of India's (RBI) 2006 Bank Authorization Policy as a quasi-natural experiment, we explore the causal relationship between bank ownership structures, local credit dynamism, and capital misallocation in underbanked districts.

The policy's implementation significantly increased private-sector bank (PVB) branches in underbanked districts. This drove an increase in credit allocated in favor of firms with

[3] Since our analysis in this part relies on bank-to-firm matching, restricting the sample to bandwidths of 10 and 15 significantly reduces the number of observations. Consequently, we are unable to perform our DiD analysis for these bandwidths in this context.

high marginal revenue product of capital (MRPK). However, the credit growth did not immediately result in proportional sales increases for these firms. Instead, it led to a temporary decline in their MRPK as firms expanded their capital bases without a corresponding rise in sales. This resulted in a decrease in capital misallocation in underbanked districts. We document that this decline in capital misallocation was short-lived, lasting two to three years. This highlights a key insight of our study: while credit growth is essential for improving capital allocation, it is equally important to allocate credit efficiently among firms. Bank ownership structure plays a central role in this process.

The policy's implementation significantly expanded private-sector bank (PVB) branches in underbanked districts, leading to greater credit allocation toward firms with high marginal revenue product of capital (MRPK). While this expansion did not immediately translate into proportional sales growth, it temporarily lowered MRPK as firms expanded their capital bases without a corresponding rise in sales. This dynamic reduced capital misallocation in underbanked districts, though the effect was short-lived, dissipating within two to three years. Beyond misallocation, the policy also influenced broader firm outcomes. Ex-ante high-MRPK firms experienced significant capital deepening—measured as higher capital per employee—and improved profit margins, yet showed no corresponding increase in gross value added. In contrast, ex-ante low-MRPK firms did not exhibit meaningful changes across these outcomes. Taken together, these results highlight two key insights: first, while credit growth is essential for reducing misallocation, its long-term effectiveness depends on allocating credit efficiently across firms; and second, bank ownership structure plays a pivotal role in shaping both the short-run and broader firm-level effects of financial sector reforms.

Our paper highlights the importance of understanding how financial intermediaries' behavior—particularly banks with differing ownership structures—shapes credit allocation. Future research could explore the reasons behind public-sector banks' static lending behavior and investigate whether more targeted reforms could encourage these institutions to adopt more efficient lending practices. Our study offers valuable insights into how financial reforms can promote more efficient resource allocation and sustainable economic growth in developing economies.

References

ALTUNBAS, Y., L. EVANS, AND P. MOLYNEUX (2001): "Bank ownership and efficiency," *Journal of Money, Credit and Banking*, 926–954.

BANERJEE, A., S. COLE, E. DUFLO, ET AL. (2004): "Banking reform in India," in *India Policy Forum*, National Council of Applied Economic Research; Brookings Institution, vol. 1, 277–332.

BAU, N. AND A. MATRAY (2023): "Misallocation and capital market integration: Evidence from India," *Econometrica*, 91, 67–106.

CALONICO, S., M. D. CATTANEO, AND M. H. FARRELL (2020): "Optimal bandwidth choice for robust bias-corrected inference in regression discontinuity designs," *The Econometrics Journal*, 23, 192–210.

CARVALHO, D. (2014): "The real effects of government-owned banks: Evidence from an emerging market," *The Journal of Finance*, 69, 577–609.

CHAKRABORTY, P. AND S. RITADHI (2022): "Do Lenders also Respond to Import Competition? Evidence from Bank-Firm Loan Level Data," *CAFRAL working paper*.

COLE, S. (2009): "Fixing market failures or fixing elections? Agricultural credit in India," *American Economic Journal: Applied Economics*, 1, 219–250.

COLEMAN, N. AND L. FELER (2015): "Bank ownership, lending, and local economic performance during the 2008–2009 financial crisis," *Journal of Monetary Economics*, 71, 50–66.

GOPINATH, G., Ş. KALEMLI-ÖZCAN, L. KARABARBOUNIS, AND C. VILLEGAS-SANCHEZ (2017): "Capital allocation and productivity in South Europe," *The Quarterly Journal of Economics*, 132, 1915–1967.

HSIEH, C.-T. AND P. J. KLENOW (2009): "Misallocation and manufacturing TFP in China and India," *The Quarterly Journal of Economics*, 124, 1403–1448.

KAPOOR, M., P. JHA, AND J. RAYCHAUDHURI (2012): "The impact of credit constraints on exporting firms: empirical evidence from India," *Available at SSRN 2023068*.

KULKARNI, N., K. MAHAJAN, AND S. RITADHI (2023): "Bank Branch Expansions and Capital Investment by Credit Constrained Firms," *Available at SSRN 4349526*.

KUMAR, N. (2020): "Political interference and crowding out in bank lending," *Journal of Financial Intermediation*, 43, 100815.

MICCO, A., U. PANIZZA, AND M. YANEZ (2007): "Bank ownership and performance. Does politics matter?" *Journal of Banking & Finance*, 31, 219–241.

RESTUCCIA, D. AND R. ROGERSON (2008): "Policy distortions and aggregate productivity with heterogeneous establishments," *Review of Economic dynamics*, 11, 707–720.

——— (2017): "The causes and costs of misallocation," *Journal of Economic Perspectives*, 31, 151–174.

SAPIENZA, P. (2004): "The effects of government ownership on bank lending," *Journal of Financial Economics*, 72, 357–384.

SRAER, D. AND D. THESMAR (2023a): "How to Use Microdata for Macro-Finance," *Annual Review of Financial Economics*, 15, 387–406.

——— (2023b): "How to use natural experiments to estimate misallocation," *American Economic Review*, 113, 906–938.

YOUNG, N. (2017): "Banking and growth: Evidence from a regression discontinuity analysis," *EBRD Working Paper*.

A Tables

Table A.1: Summary Statistics for Underbanked and Banked Districts

	All		Underbanked	Banked	
	Mean	SD	Mean	Mean	Difference in Mean (4-3)
Variable	(1)	(2)	(3)	(4)	(5)
Credit Growth (in %)	19.9	18.8	19.3	20.1	0.8
Deposits Growth (in %)	13.4	11.5	13.2	13.8	0.6
Public Branch Share (in %)	65.9	19.1	62.5	71.4	8.9***
Private Branch Share (in %)	10.5	12.1	7.2	13.4	6.2***
Branches per Million Population	83.4	376	48.6	144.6	96***
New Branches	0.42	7.1	-0.3	1.9	2.2***
Observations	2,895		1,849	1,046	

Note: The summary statistics show the mean value of various indicators at the district level for the period of 2001 to 2005, the period prior to BAP. Columns 1 and 2 capture the mean and standard deviation of the complete set of districts pre-period, whereas columns 3 and 4 capture the mean value for banked and underbanked districts. Column 5 captures the difference in the mean by district type and reports the level of significance of the difference. New branches are the net of new branches that are operational over the previous year. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.2: Summary Statistics of Loan Amounts by Bank Ownership

Creditor Type	Mean	SD	P25	P50	P75	Min	Max	Count
	Panel A: Firm-Bank Loan Data							
Private-Sector Banks	4.53	145.27	0.08	0.59	2.20	0.00	22,755.55	25,847
Public-Sector Banks	6.68	295.82	0.17	0.86	3.27	0.00	62,211.63	71,324
Total	6.11	264.29	0.14	0.78	2.92	0.00	62,211.63	97,171
	Panel B: Merged Firm-Bank Loan Data and Prowess							
Private-Sector Banks	3.37	15.71	0.09	0.70	2.35	0.00	562.70	8,745
Public-Sector Banks	5.78	29.69	0.29	1.17	3.72	0.00	1,500.00	26,191
Total	5.18	26.90	0.23	1.02	3.39	0.00	1,500.00	34,936

Note: Panel A documents the summary statistics of loan amounts by creditor type (private-sector banks and public-sector banks) using the whole sample of firm-bank loan matched data for the period of 2001-2011. Panel B presents the summary statistics of loan amounts by creditor type (private-sector banks and public-sector banks) using the merged sample of firm-bank loan matched data and CMIE Prowess for the period 2001-2011, focusing only on manufacturing sector firms. Loan amounts are denoted in INR millions and deflated using the 2005 GDP deflator.

Table A.3: Summary Statistics of Firm Borrower-level Variables by Bank Ownership

Variable	Mean	SD	P50	Min	Max	Count
	Panel A: Firms with Private-Sector Banks					
Log MRPK	2.91	1.29	2.91	0.00	10.58	4,402
Log Assets	2.86	1.58	2.75	0.00	9.87	4,601
Log Sales	2.77	1.49	2.73	0.00	9.35	4,404
Log Cash Profits	0.99	1.07	0.69	-4.56	7.47	4,447
	Panel B: Firms with Public-Sector Banks					
Log MRPK	2.70	1.24	2.70	0.00	12.00	11,401
Log Assets	2.42	1.53	2.21	0.00	9.87	12,149
Log Sales	2.39	1.45	2.27	0.00	9.38	11,407
Log Cash Profits	0.73	0.98	0.38	-4.85	7.67	11,580
	Panel C: Total					
Log MRPK	2.76	1.26	2.76	0.00	12.00	15,803
Log Assets	2.54	1.56	2.35	0.00	9.87	16,750
Log Sales	2.50	1.47	2.40	0.00	9.38	15,811
Log Cash Profits	0.80	1.01	0.46	-4.85	7.67	16,027

Note: This table presents the summary statistics of selected firm-level variables (logarithms of MRPK, assets, sales, and cash profits) categorized by creditor type (public-sector and private-sector banks) for the period of 2001-2011. The variables for assets, sales, and cash profits are deflated using the 2005 GDP deflator.

Table A.4: Summary Statistics of Firm-level Log MRPK in Banked and Underbanked Districts

Variable	Mean	SD	P50	Min	Max	Count
	Panel A: 2001-2011 Period					
Log MRPK in Banked Districts	2.36	1.43	2.31	0.00	12.00	41,698
Log MRPK in Underbanked Districts	2.10	1.25	2.06	0.00	10.52	8,664
Total	2.32	1.40	2.26	0.00	12.00	50,362
	Panel B: 2001-2005 Period					
Log MRPK in Banked Districts	2.21	1.35	2.15	0.00	10.58	15,187
Log MRPK in Underbanked Districts	1.92	1.14	1.89	0.00	6.36	3,066
Total	2.16	1.32	2.10	0.00	10.58	18,253
	Panel C: 2006-2011 Period					
Log MRPK in Banked Districts	2.45	1.47	2.40	0.00	12.00	26,511
Log MRPK in Underbanked Districts	2.20	1.29	2.16	0.00	10.52	5,598
Total	2.41	1.44	2.36	0.00	12.00	32,109

Note: This table presents the summary statistics of the logarithm of firm-level MRPK for underbanked and banked districts before and after the bank branch expansion policy implemented in the fiscal year of 2006. Firm-level MRPK is calculated as the firm-level total sales per total capital. The table uses data from the 2001-2011 period.

Table A.5: Difference-in-Differences Regression Result for Bank Branch Expansion by Bank Ownership, No Bandwidth

	Total Branches			PSB Branches			PVB Branches		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Underbanked × Post	-0.025**	-0.017	-0.013	-0.027**	-0.025*	-0.023*	0.005	0.095**	0.112***
	(0.012)	(0.015)	(0.014)	(0.011)	(0.014)	(0.013)	(0.045)	(0.042)	(0.042)
Distirct FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State×Year FE	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
District Trends	No	No	Yes	No	No	Yes	No	No	Yes
R^2	0.993	0.994	0.994	0.993	0.994	0.995	0.949	0.970	0.971
Observations	6,387	6,332	6,116	6,387	6,332	6,116	3,920	3,835	3,746

Note: The table above is a Difference-in-Differences regression for the log of number of bank branches as specified in equation (1) to assess the effect of policy change on refined underbanked districts vis à vis banked districts. Here, we are only looking at the average treatment effect of the treated on overall banks and, then, separately by bank ownership. The dependent variables are the log of total number of branches at the district level. It is regressed against $Post_t \times Underbanked_d$ and $Underbanked_d$ takes the value of 1 if the district is underbanked, whereas $Post_t$ takes a value of 1 if the year is on and after 2006. The first columns for each variable refer to the results with limited fixed effects, while the last column constitutes all the fixed effects and other controls. The study spans data from fiscal year 2001 to 2011. The analysis is at the district level, is annual, and covers all Indian districts as per Census 2001 in an unbalanced panel. Standard errors (in parentheses) are clustered at the district level. $^* p < 0.10$, $^{**} p < 0.05$, $^{***} p < 0.01$.

Table A.6: Difference-in-Differences Regression Result for Capital Misallocation, No Bandwidth

	(1) Log MRPK	(2) Log MRPK	(3) Log MRPK	(4) Log MRPK	(5) Log MRPK	(6) Log MRPK	(7) Log MRPK	(8) Log MRPK
Underbanked × Post	0.0867	0.0870	0.0689	-0.0643	-0.00230	-0.00150	-0.00371	-0.0415
	(0.0628)	(0.0624)	(0.0607)	(0.0621)	(0.0543)	(0.0536)	(0.0534)	(0.0579)
Underbanked × Post × $I^{HighMRPK}$					-0.736***	-0.733***	-0.763***	-0.774***
					(0.155)	(0.154)	(0.145)	(0.160)
District FE	No	No	Yes	Yes	No	No	Yes	Yes
Year FE	No	Yes	Yes	Yes	No	Yes	Yes	Yes
State×Year FE	No	No	No	Yes	No	No	No	Yes
District Trends	No	No	No	Yes	No	No	No	Yes
R^2	0.00758	0.0107	0.0628	0.0628	0.256	0.260	0.301	0.306
Observations	50,293	50,293	50,285	45,571	41,618	41,618	41,616	37,953

Note: This table presents the difference-in-differences estimates of the regression model outlined in equation (3). Post_t indicates whether time t is post-treatment (i.e., after the year 2006), with a value of 1 for post-treatment periods and 0 otherwise. Underbanked_d is a dummy variable indicating whether district d is underbanked (1 for underbanked districts, 0 otherwise). I^{HighMRPK} is a dummy variable set to 1 if a firm's ex-ante MRPK before the policy exceeds the industry mean MRPK within the year. District-level control variables are the logarithm of the population and the logarithm of the number of bank branches, which interact with a linear time trend. Here, we are looking at the average treatment effect on the all treated districts compared with the control districts. Standard errors (in parentheses) are clustered at the district level. $^* p < 0.05$, $^{**} p < 0.01$, $^{***} p < 0.001$.

Table A.7: Difference-in-Differences Regression Result for Capital Misallocation, Bandwidth 15

	(1) Log MRPK	(2) Log MRPK	(3) Log MRPK	(4) Log MRPK	(5) Log MRPK	(6) Log MRPK	(7) Log MRPK	(8) Log MRPK
Underbanked × Post	0.212*	0.210*	0.141	0.0397	0.0840	0.0823	0.0615	-0.0267
	(0.102)	(0.103)	(0.0939)	(0.114)	(0.0794)	(0.0777)	(0.0764)	(0.103)
Underbanked × Post × $I^{HighMRPK}$					-0.756***	-0.754***	-0.699***	-0.634**
					(0.212)	(0.212)	(0.185)	(0.201)
District FE	No	No	Yes	Yes	No	No	Yes	Yes
Year FE	No	Yes	Yes	Yes	No	Yes	Yes	Yes
State×Year FE	No	No	No	Yes	No	No	No	Yes
District Trends	No	No	No	Yes	No	No	No	Yes
R^2	0.00929	0.0141	0.112	0.114	0.234	0.240	0.315	0.316
Observations	11,539	11,539	11,538	10,907	9,411	9,411	9,411	8,968

Note: This table presents the difference-in-differences estimates of the regression model outlined in equation (3). Post_t indicates whether time t is post-treatment (i.e., after the year 2006), with a value of 1 for post-treatment periods and 0 otherwise. Underbanked_d is a dummy variable indicating whether district d is underbanked (1 for underbanked districts, 0 otherwise). I^{HighMRPK} is a dummy variable set to 1 if a firm's ex-ante MRPK before the policy exceeds the industry mean MRPK within the year. District-level control variables are the logarithm of the population and the logarithm of the number of bank branches, which interact with a linear time trend. Here, we are looking at the average treatment effect on the treated districts that lie within the bandwidth of *15 bank branches per million* around the national average threshold and compared with the control districts. Standard errors (in parentheses) are clustered at the district level. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

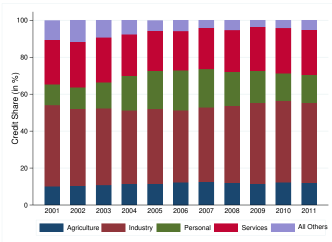
Table A.8: Difference-in-Differences Regression Result for Capital Misallocation, Bandwidth 10

	(1) Log MRPK	(2) Log MRPK	(3) Log MRPK	(4) Log MRPK	(5) Log MRPK	(6) Log MRPK	(7) Log MRPK	(8) Log MRPK
Underbanked × Post	0.242	0.245	0.144	0.00572	0.0502	0.0541	0.0121	-0.0705
	(0.134)	(0.134)	(0.125)	(0.129)	(0.0963)	(0.0933)	(0.0917)	(0.128)
Underbanked × Post × $I^{HighMRPK}$					-0.803**	-0.799**	-0.696***	-0.612**
					(0.244)	(0.243)	(0.192)	(0.206)
District FE	No	No	Yes	Yes	No	No	Yes	Yes
Year FE	No	Yes	Yes	Yes	No	Yes	Yes	Yes
State×Year FE	No	No	No	Yes	No	No	No	Yes
District Trends	No	No	No	Yes	No	No	No	Yes
R^2	0.0104	0.0156	0.107	0.112	0.235	0.242	0.308	0.308
Observations	8,335	8,335	8,334	7,843	6,882	6,882	6,882	6,525

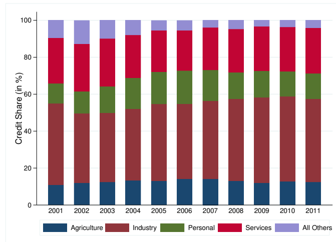
Note: This table presents the difference-in-differences estimates of the regression model outlined in equation (3). Post_t indicates whether time t is post-treatment (i.e., after the year 2006), with a value of 1 for post-treatment periods and 0 otherwise. Underbanked_d is a dummy variable indicating whether district d is underbanked (1 for underbanked districts, 0 otherwise). I^{HighMRPK} is a dummy variable set to 1 if a firm's ex-ante MRPK before the policy exceeds the industry mean MRPK within the year. District-level control variables are the logarithm of the population and the logarithm of the number of bank branches, which interact with a linear time trend. Here, we are looking at the average treatment effect on the treated districts that lie within the bandwidth of *10 bank branches per million* around the national average threshold and compared with the control districts. Standard errors (in parentheses) are clustered at the district level. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

B Figures

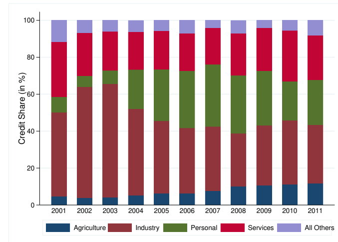
Figure B.1: Credit Share by Loan and Bank Ownership



(a) All Banks



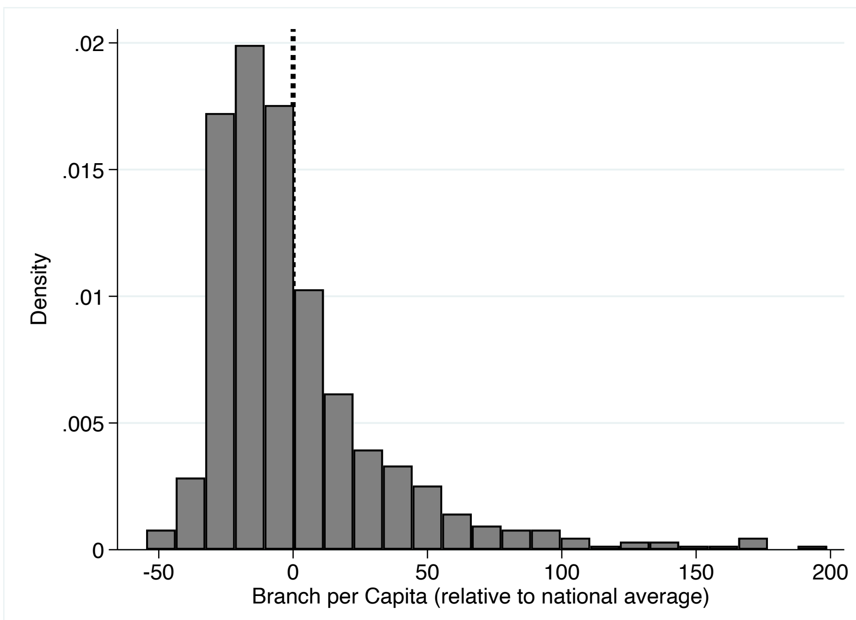
(b) Public-Sector Banks



(c) Private-Sector Banks

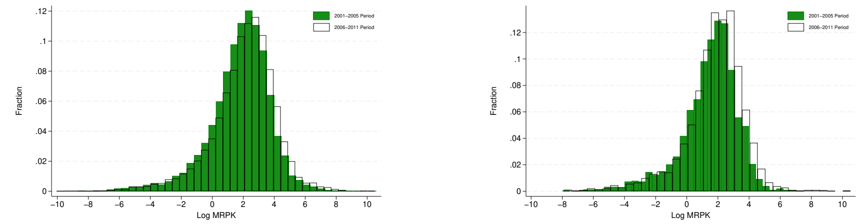
Note: The figure shows the overall distribution of various types of loans over time and the breakup of those loans by bank ownership. Industry loans comprise the majority of loans issued for the banking industry but are declining, with a substitution towards services and personal loans.

Figure B.2: Distribution of Districts around the National Average of Branch Density



Note: The figure shows the distribution of the difference of branch density, defined at the district level as $Runvar_d = Branch\ Per\ Capita_d - \overline{Branch\ Per\ Capita}$. Where $Branch\ Per\ Capita_d$ refers to the bank branch density in district d in 2005, while $\overline{Branch\ Per\ Capita}$ is the national average bank branch density in 2005. Districts are classified as "underbanked" if $Runvar_d < 0$.

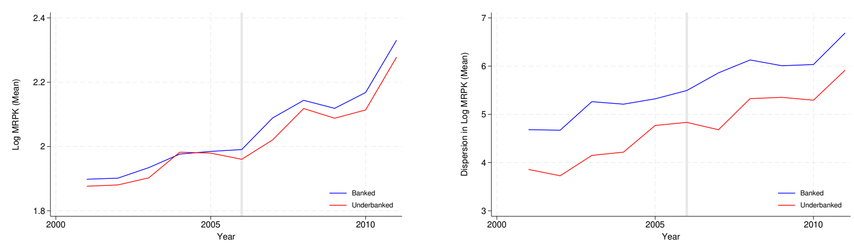
Figure B.3: Histogram of Firm-level Log MRPK in Banked and Underbanked Districts



(a) Banked Districts (b) Underbanked Districts

Note: This figure shows the histogram of the logarithm of firm-level MRPK for underbanked and banked districts before and after the bank branch expansion policy implemented in the fiscal year of 2006. Firm-level MRPK is calculated as the firm-level total sales per total capital. The figure uses data from the 2001-2011 period.

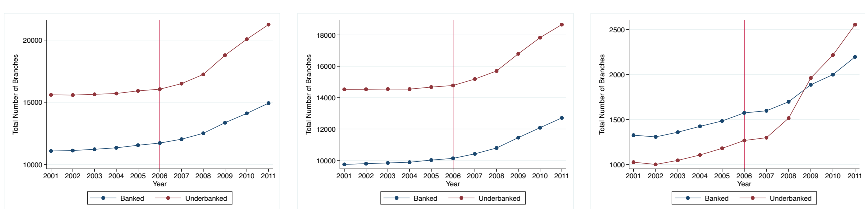
Figure B.4: Log MRPK and Dispersion in Log MRPK over Time in Underbanked and Banked Districts



(a) Log MRPK (b) Dispersion in Log MRPK

Note: Figure (a) presents the annual mean of the logarithm of firm-level MRPK in underbanked and banked districts, and Figure (b) presents the annual mean of dispersion in the logarithm of firm-level MRPK (a measure of capital misallocation) in underbanked and banked districts. MRPK is calculated as the firm-level total sales divided by total capital. The figure uses data from the 2001-2011 period. The grey bar in 2006 indicates the fiscal year of implementing the bank authorization policy.

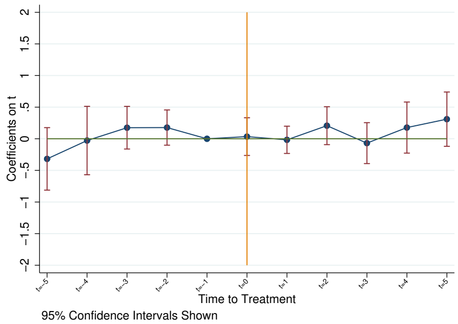
Figure B.5: Trends in the Number of Bank Branches by Bank Ownership



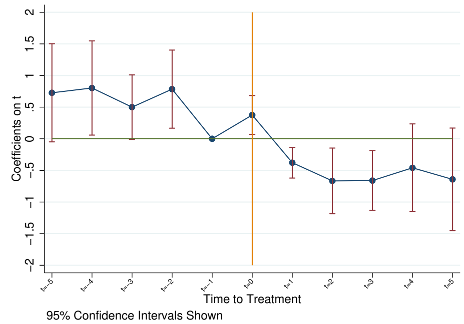
(a) All Banks (b) Public Banks (c) Private Banks

Note: The figure shows the trend of evolution of the number of branches by different types of banks over time for underbanked and banked districts that lie within the bandwidth of *20 bank branches per million* around the national average threshold. The figure presents trends for all banks and by ownership of the banks. The sample period spans from 2001 to 2011, with 2006 as the fiscal year implementing the bank branch expansion policy.

Figure B.6: Dynamic DiD - Single District Firms and Refined Districts



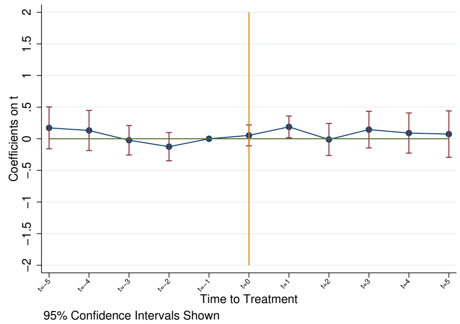
(a) Ex ante Low MRPK Firms



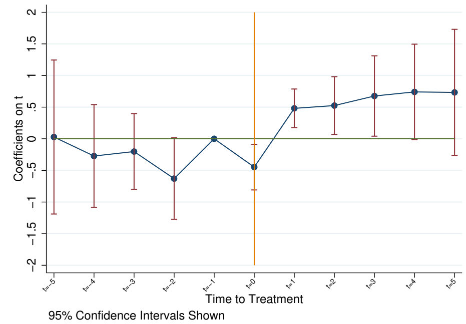
(b) Ex ante High MRPK Firms

Note: This figure presents the dynamic difference-in-differences estimates of the regression model outlined in equation (4) for single district firms and refined districts where we keep underbanked districts that experienced at least one bank branch opening following the policy. We categorize firms as ex-ante high (low) MRPK if their average MRPK is above (below) the 2-digit industry-level mean from 2001-2005. The regression model includes district-level control variables (the logarithm of the population and the logarithm of the number of bank branches interacted with a linear time trend), year, district, and state-year fixed effects. Here, we are looking at the average treatment effect on the treated districts that lie within the bandwidth of *20 bank branches per million* around the national average threshold and compared with the control districts. Standard errors are clustered at the district level. The sample period spans from 2001 to 2011, with 2006 as the fiscal year of $t = 0$, corresponding to the implementation of the bank branch expansion policy. Confidence intervals are calculated at the 95% level.

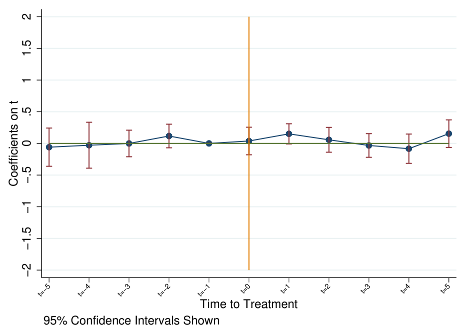
Figure B.7: Dynamic DiD for Firm Outcomes - Single District Firms and Refined Districts



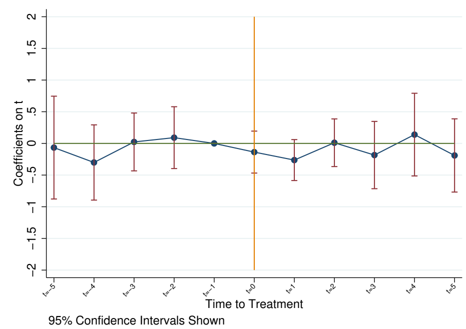
(a) Ex ante Low MRPK Firms: Log Capital/Worker



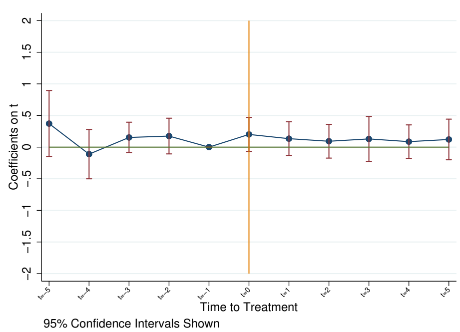
(b) Ex ante High MRPK Firms: Log Capital/Worker



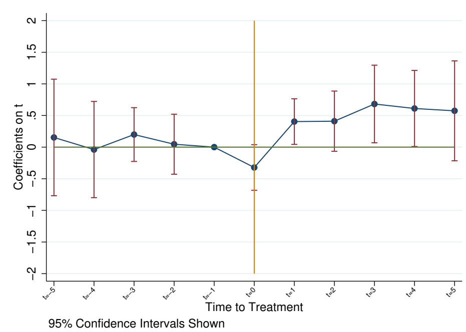
(c) Ex ante Low MRPK Firms: Log GVA



(d) Ex ante High MRPK Firms: Log GVA



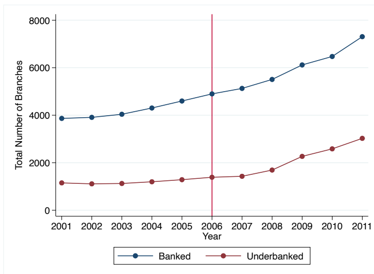
(e) Ex ante Low MRPK Firms: Profit Margin



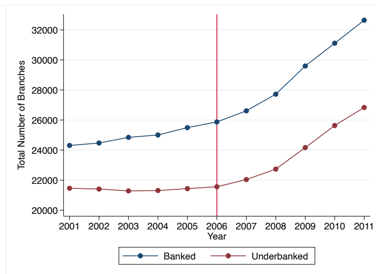
(f) Ex ante High MRPK Firms: Profit Margin

Note: This figure presents the dynamic difference-in-differences estimates from the regression model outlined in equation (4) for firm-level outcomes: capital per employee (a proxy for capital deepening), gross value added, and profit margin. The analysis is restricted to single-district firms and refined underbanked districts that experienced at least one bank branch opening following the policy. We categorize firms as ex-ante high (low) MRPK if their average MRPK is above (below) the 2-digit industry-level mean from 2001-2005. The regression model includes district-level control variables (the logarithm of the population and the logarithm of the number of bank branches interacted with a linear time trend), year, district, and state-year fixed effects. Here, we are looking at the average treatment effect on the treated districts that lie within the bandwidth of *20 bank branches per million* around the national average threshold and compared with the control districts. Standard errors are clustered at the district level. The sample period spans from 2001 to 2011, with 2006 as the fiscal year of $t = 0$, corresponding to the implementation of the bank branch expansion policy. Confidence intervals are calculated at the 95% level.

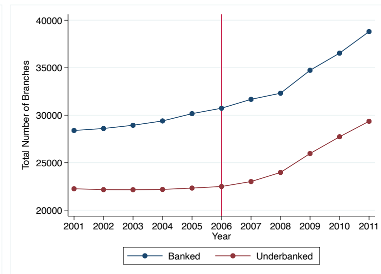
Figure B.8: Trends in Number of Bank Branches by Bank Ownership, Different Band-widths



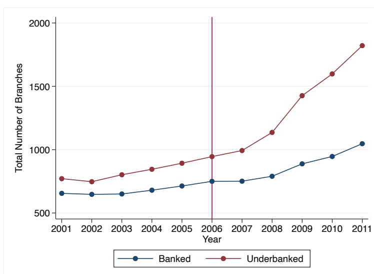
(a) Private Banks - No BW



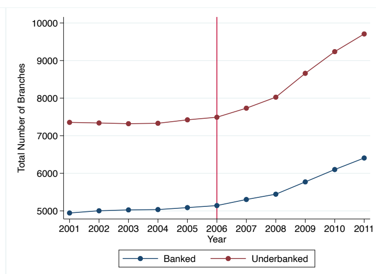
(b) Public Banks - No BW



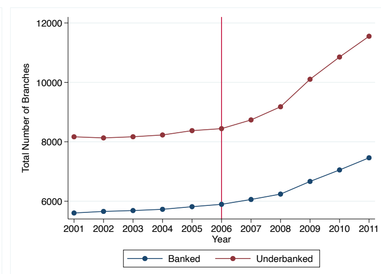
(c) All Banks - No BW



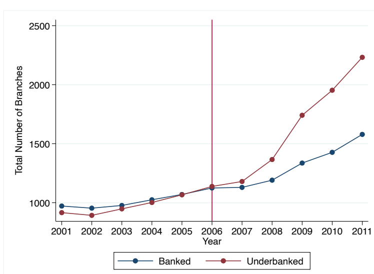
(d) Private Banks - BW10



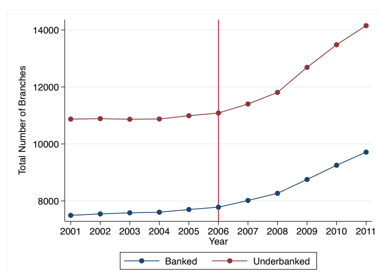
(e) Public Banks - BW10



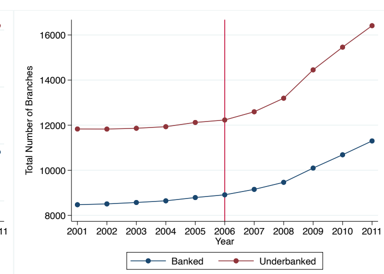
(f) All Banks - BW10



(g) Private Banks - BW15



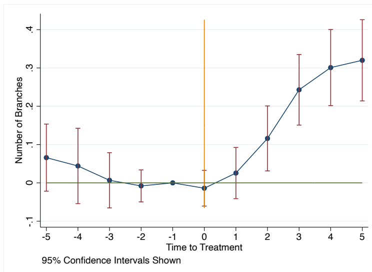
(h) Public Banks - BW15



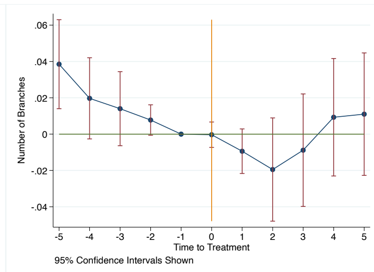
(i) All Banks - BW15

Note: This figure presents the trend of evolution of the number of branches by different types of banks over time and across various band-width around the cut-off for underbanked and non underbanked districts. The regression model includes district-level control variables (the logarithm of the population and the logarithm of the number of bank branches which are interacted with a linear time trend), year, district, and state-year fixed effects. Standard errors are clustered at the district level. The sample period spans from 2001 to 2011, with 2006 as the fiscal year of $t = 0$, corresponding to the implementation of the bank branch expansion policy. Confidence intervals are calculated at the 95% level.

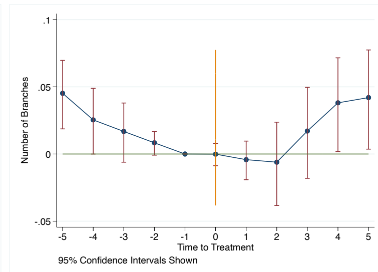
Figure B.9: Dynamic DiD - Heterogeneous Bank-level Responses, Different Bandwidths



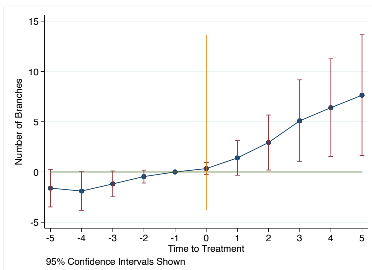
(a) Private Banks - No BW



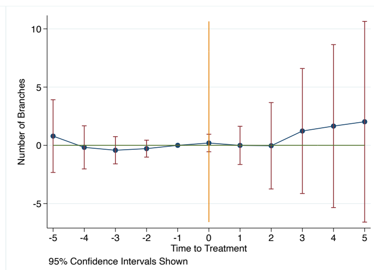
(b) Public Banks - No BW



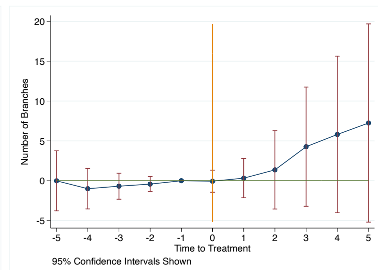
(c) All Banks - No BW



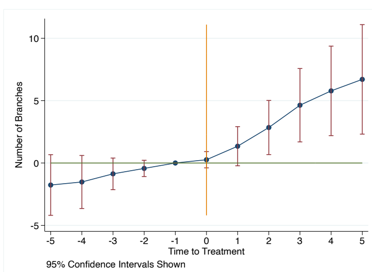
(d) Private Banks - BW 10



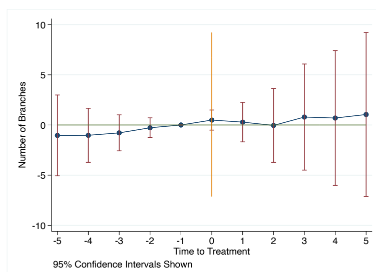
(e) Public Banks - BW 10



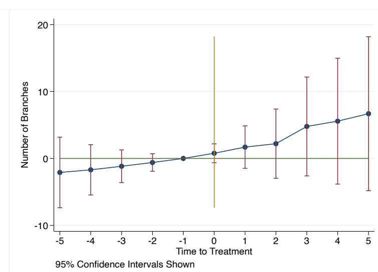
(f) All Banks - BW 10



(g) Private Banks - BW 15



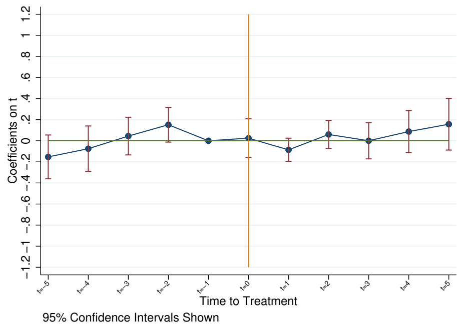
(h) Public Banks - BW 15



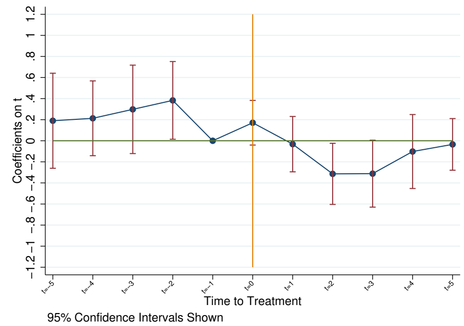
(i) All Banks - BW 15

Note: This figure shows the dynamic evolution of number of branches by different types of bank over time and across various band-width around the cut-off for underbanked and non underbanked districts. The regression model includes district-level control variables (the logarithm of the population and the logarithm of the number of bank branches which are interacted with a linear time trend), year, district, and state-year fixed effects. Standard errors are clustered at the district level. The sample period spans from 2001 to 2011, with 2006 as the fiscal year of $t = 0$, corresponding to the implementation of the bank branch expansion policy. Confidence intervals are calculated at the 95% level.

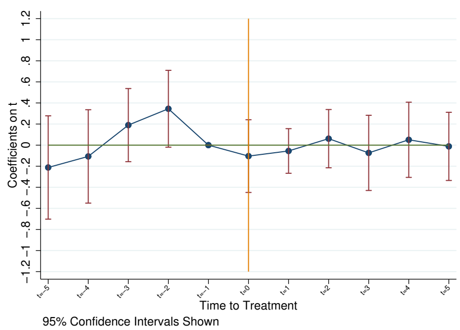
Figure B.10: Dynamic DiD - Capital Misallocation, Different Bandwidths



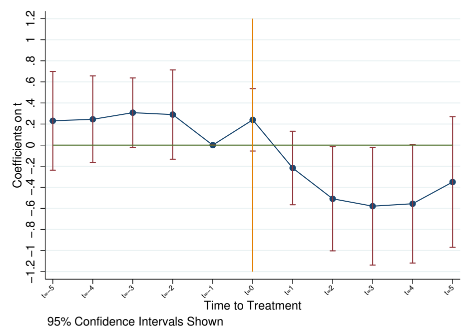
(a) Ex ante Low MRPK - No BW



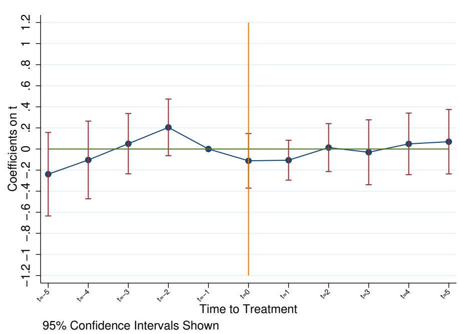
(b) Ex ante High MRPK - No BW



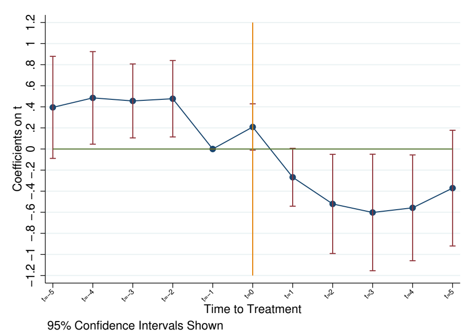
(c) Ex ante Low MRPK - BW10



(d) Ex ante High MRPK - BW10



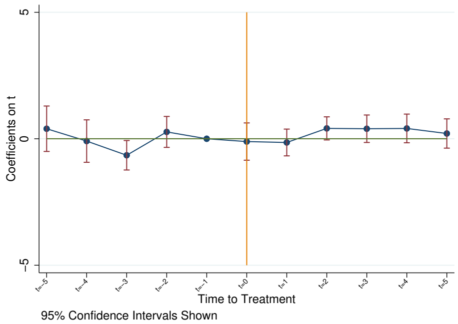
(e) Ex ante Low MRPK - BW15



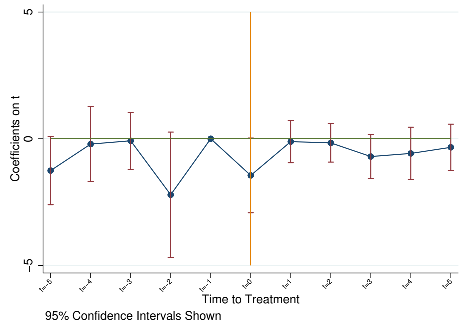
(f) Ex ante High MRPK - BW15

Note: The figure shows the benchmark dynamic DiD results on capital misallocation by different bandwidths. We categorize firms as ex-ante high (low) MRPK if their average MRPK is above (below) the 2-digit industry-level mean from 2001-2005. The regression model includes district-level control variables (the logarithm of the population and the logarithm of the number of bank branches which are interacted with a linear time trend), year, district, and state-year fixed effects. Standard errors are clustered at the district level. The sample period spans from 2001 to 2011, with 2006 as the fiscal year of $t = 0$, corresponding to the implementation of the bank branch expansion policy. Confidence intervals are calculated at the 95% level.

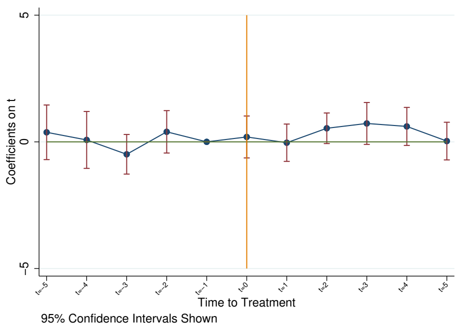
Figure B.11: Dynamic DiD - Public and Private Sector Bank Loans, No Bandwidth



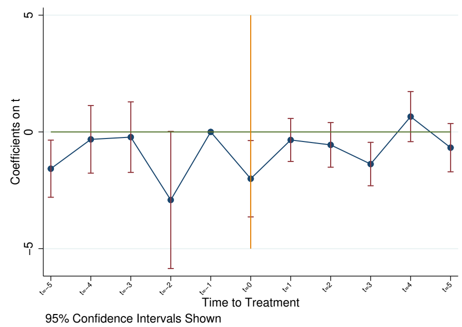
(a) All - PSB Loans



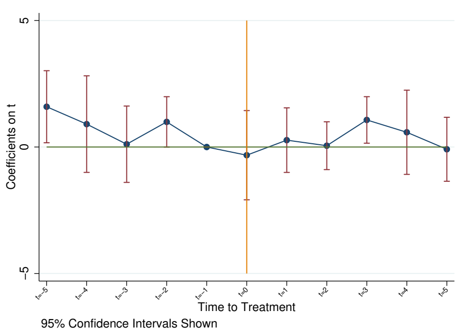
(b) All - PVB Loans



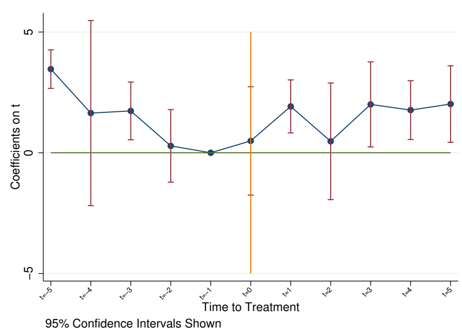
(c) Ex ante Low MRPK - PSB Loans



(d) Ex ante Low MRPK - PVB Loans



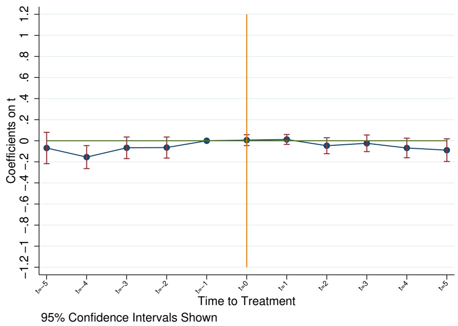
(e) Ex ante High MRPK - PSB Loans



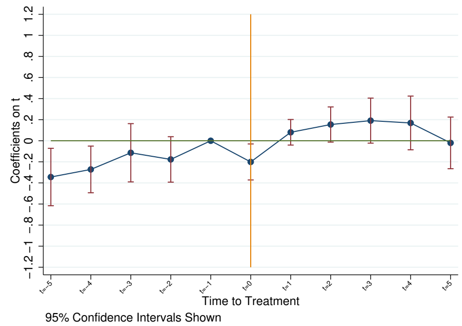
(f) Ex ante High MRPK - PVB Loans

Note: The figure shows the benchmark dynamic DiD results on public and private sector bank loans by different bandwidths. We categorize firms as ex-ante high (low) MRPK if their average MRPK is above (below) the 2-digit industry-level mean from 2001-2005. The regression model includes district-level control variables (the logarithm of the population and the logarithm of the number of bank branches which are interacted with a linear time trend), year, district, and state-year fixed effects. Loan amounts are deflated. Standard errors are clustered at the district level. The sample period spans from 2001 to 2011, with 2006 as the fiscal year of $t = 0$, corresponding to the implementation of the bank branch expansion policy. Confidence intervals are calculated at the 95% level.

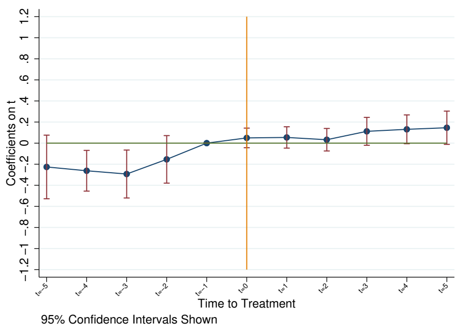
Figure B.12: Dynamic DiD - Firm Capital, Different Bandwidths



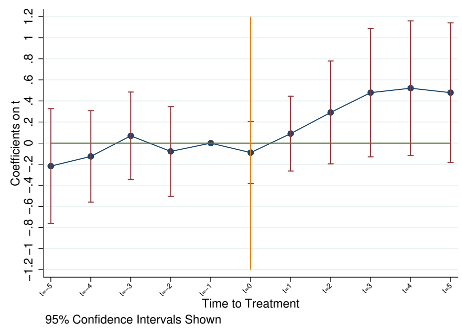
(a) Ex ante Low MRPK - No BW



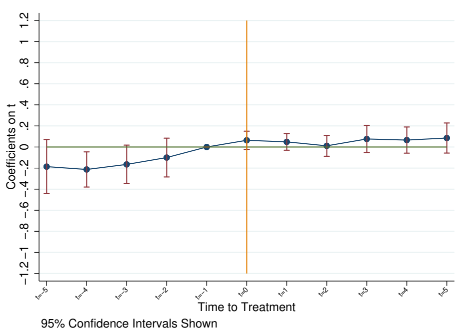
(b) Ex ante High MRPK - No BW



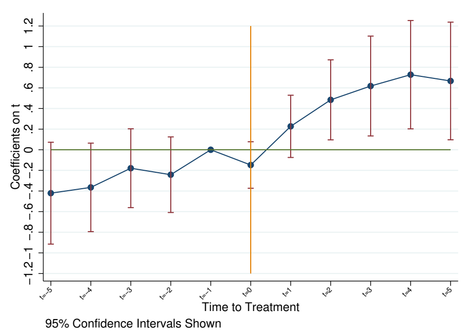
(c) Ex ante Low MRPK - BW10



(d) Ex ante High MRPK - BW10



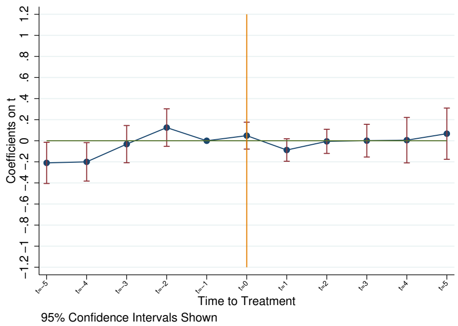
(e) Ex ante Low MRPK - BW15



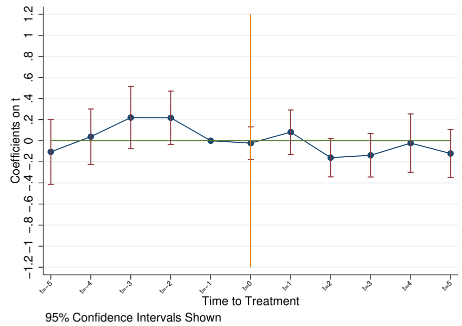
(f) Ex ante High MRPK - BW15

Note: The figure shows the benchmark dynamic DiD results on firm capital by different bandwidths. We categorize firms as ex-ante high (low) MRPK if their average MRPK is above (below) the 2-digit industry-level mean from 2001-2005. The regression model includes district-level control variables (the logarithm of the population and the logarithm of the number of bank branches which are interacted with a linear time trend), year, district, and state-year fixed effects. Standard errors are clustered at the district level. The sample period spans from 2001 to 2011, with 2006 as the fiscal year of $t = 0$, corresponding to the implementation of the bank branch expansion policy. Confidence intervals are calculated at the 95% level.

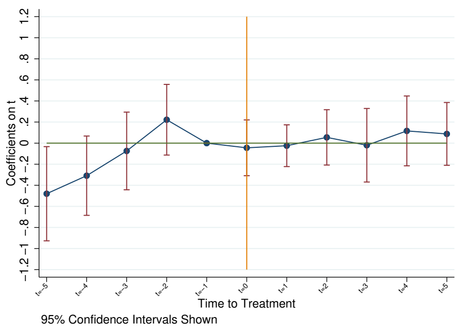
Figure B.13: Dynamic DiD - Firm Sales, Different Bandwidths



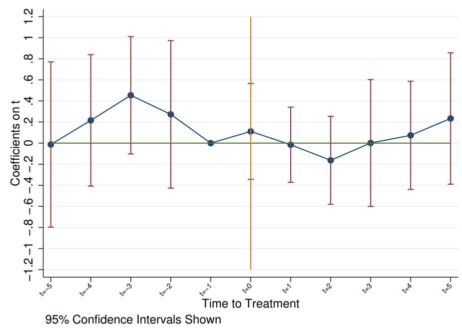
(a) Ex ante Low MRPK - No BW



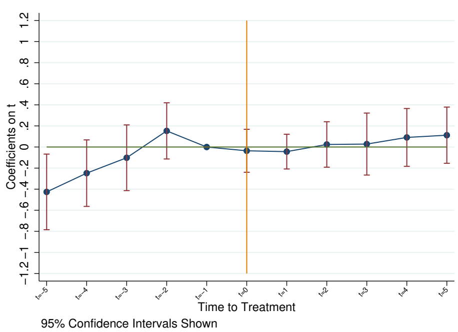
(b) Ex ante High MRPK - No BW



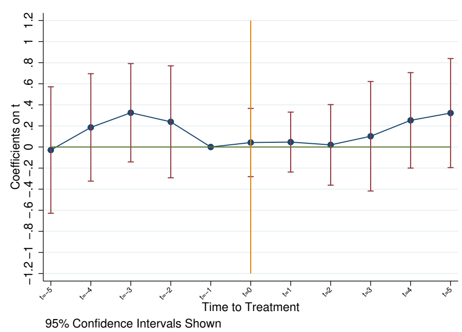
(c) Ex ante Low MRPK - BW10



(d) Ex ante High MRPK - BW10



(e) Ex ante Low MRPK - BW15



(f) Ex ante High MRPK - BW15

Note: The figure shows the benchmark dynamic DiD results on firm sales by different bandwidths. We categorize firms as ex-ante high (low) MRPK if their average MRPK is above (below) the 2-digit industry-level mean from 2001-2005. The regression model includes district-level control variables (the logarithm of the population and the logarithm of the number of bank branches which are interacted with a linear time trend), year, district, and state-year fixed effects. Standard errors are clustered at the district level. The sample period spans from 2001 to 2011, with 2006 as the fiscal year of $t = 0$, corresponding to the implementation of the bank branch expansion policy. Confidence intervals are calculated at the 95% level.