Bank Expansion, Firm Dynamics, and Structural Transformation: Evidence from India's Policy Experiment[†]

Bharat Chandar [‡]
Dian Jiao [§]
Marshall Mo [¶]

October 27, 2024

Abstract

This paper examines the impacts of bank expansion on firm dynamics and labor allocation, exploiting a policy experiment in India designed to encourage bank expansion in "under-banked" districts. Empirical findings demonstrate significant growth in manufacturing firms in these districts due to eased credit access, resulting in increased capital accumulation, sales revenue, and employment. However, the expansion predominantly benefited incumbent firms, with minimal stimulation of firm entry or product innovation. The reform also induced notable labor reallocation towards manufacturing sectors, particularly in areas with lower agricultural productivity.

[†] We would like to express our gratitude to Pete Klenow, Eric Verhoogen, and Jack Willis for their invaluable guidance and support throughout the project. We are also thankful to Melanie Morten, Kim Fe Cramer, Giorgia Barboni, and the participants in the Development Colloquium at Columbia University, Stanford history group meeting, and Webinar series in Finance and Development for their helpful comments and insights. Financial support from STEG is also acknowledged. Special thanks to Winston Xu for his significant contribution to the early stage of the project. All errors are our own.

[‡] Stanford University; chandarb@stanford.edu

[§] Columbia University; dj2531@columbia.edu

[¶] Stanford University; erxiaomo@stanford.edu

1 Introduction

Access to financial services remains markedly limited in developing countries. An estimate from the World Bank posits that approximately 1.7 billion people, primarily residing in underdeveloped countries, are devoid of basic banking facilities, such as checking accounts or standard savings products (Demirguc-Kunt et al., 2018). This situation starkly contrasts with developed economies like the United States, where a mere 4.5% of households are "unbanked" (FDIC, 2021). This deprivation from finance could impede economic development by curbing aggregate investment and employment and distorting capital allocation among firms and potential entrepreneurs (Hsieh and Klenow, 2009; Buera et al., 2011; Bazzi et al., 2021; Fonseca and Matray, 2022).

Many developing countries have implemented place-based policies to stimulate bank expansion in lagging regions, believing that these policies have the potential to ignite economic growth and alleviate poverty through better access to financial services. Previous literature, utilizing both Randomized Controlled Trials (RCTs) and natural experiments, tends to highlight substantial benefits to industrial growth and household well-being in areas that have seen an influx of banking services (Burgess and Pande, 2005; Bruhn and Love, 2014; Young, 2017; Cramer, 2021; Fonseca and Matray, 2022; Barboni et al., 2021). However, less is understood about the costs and aggregate welfare effects of these programs.[1] It is conceivable that regions under-served by banks could have an inherently low demand for credit, possibly due to low productivity levels or other market frictions.[2] Under such conditions, expanding the presence of bank branches in places with low credit demand could be counterproductive and even lead to greater misallocation of resources.

In this study, we investigate the impacts of bank expansion on firm dynamics and labor allocation, employing both empirical and theoretical approaches. Our analysis takes advantage of a nationwide policy experiment in India, which was implemented by the Reserve Bank of India in 2005. The purpose of this policy was to encourage the opening of branches in "under-banked" districts—regions where the population-to-branch ratio surpassed the national average. This arbitrary policy cutoff presents us with an ideal setting to estimate the causal effects of bank expansion using a regression discontinuity design (RDD), wherein we compare districts that barely exceed or fall below the criterion for being classified as under-

[1]One noteworthy exception is Ji et al. (2023), which integrates empirical evidence of branch openings in local markets with a spatial general equilibrium model to quantify the aggregate effects of bank expansion.

[2]Previous studies have shed light on several industrial policies that could potentially contribute to such market frictions. These policies include priority sector lending (Banerjee and Duflo, 2014), restrictions on Foreign Direct Investment (FDI) (Bau and Matray, 2023), contractual frictions (Bertrand et al., 2021), and reservations for Small-Scale Industries (Martin et al., 2017). In section 4, we outline the theoretical framework for how such market frictions can affect firm dynamics.

banked. The banking industry responded strongly to the policy, leading to an increase in the number of branches, deposits, and credit in those under-banked districts, thereby creating exogenous shocks to financial access.

Leveraging plant-level data from the Annual Survey of Industries (ASI) database from 1998 to 2013, we document a significant expansion of manufacturing firms in under-banked districts post-reform. Our findings indicate that the reform eased firms' access to credit, stimulating capital accumulation and driving growth in sales revenue and employment. However, this growth was exclusively fueled by incumbent firms that employed more resources to increase their production of existing products. Contrary to findings in prior studies (Kerr and Nanda, 2009; Bazzi et al., 2021; Fonseca and Matray, 2022), the reform failed to spur greater firm dynamics. Although treated districts saw an uptick in the exit rate for small firms (with less than 20 employees), the entry rate remained unchanged post-reform. We similarly report that the reform had limited effects on product innovation and creative destruction. In line with the lack of enhanced firm dynamics and product innovation, firms in the treated districts did not witness an improvement in their Total Factor Productivity (TFP).

Treating each Indian district as a distinct local market, we delve into the reform's equilibrium impacts on structural transformation by leveraging the village-level population census data.[3] We observe that bank expansion induced significant labor reallocation towards the manufacturing sector, primarily by creating more manufacturing jobs. This effect is particularly pronounced in villages with lower agricultural productivity, implying that the reform potentially mitigates labor misallocation across different sectors.

We are developing a quantitative model to study redistributive and welfare consequences of counterfactual banking policies. The quantitative model features heterogeneous firms that require financing to operate. Small firms face high interest rates, so they choose alternative funding sources instead of banks. Big firms face lower interest rates, so they benefit and grow more following bank expansion. In equilibrium this leads to expansion of the whole market but potentially crowd out of smaller firms. These predictions match our empirical findings regarding firm dynamics. We use our model to study the welfare effects of the government's policy to encourage banks to enter underbanked areas. In our model the key tradeoff from the policy is redistributive. In order to enter a growing, profitable district that

[3]Treating each district as a distinct local labor market is consistent with evidence of low levels of migration in our study context. Existing literature has highlighted substantial labor mobility costs in India, potentially confining labor to agriculture even in the face of low productivity. Munshi and Rosenzweig (2016) highlights that the wage gap between urban and rural areas in India is substantial, with urban wages being over 47% higher than rural wages for less educated workers engaged in menial tasks. Using detailed district-wise migration flow data from the 2001 population census, Kone et al. (2018) underscores that the 5-year inter-district migration rate in India is remarkably low, at a mere 2.8%. In stark contrast, the 5-year inter-prefecture migration rate in China stands significantly higher at 10%.

already has above-median bank access, a bank must also enter an underbanked district. This distorts the choices for where banks enter, but could lead to benefits to districts with poor existing financial institutions. We study these heterogeneous effects across districts. We then consider alternative banking policies, such as direct taxes and subsidies to enter in different districts. We are extending the model to speak to our empirical results about structural transformation.

This paper connects most closely to three main streams of literature. First, we add to a growing body of literature studying the impacts of quasi-experimental and experimental credit shocks in developing countries (Bazzi et al., 2021; Breza and Kinnan, 2021; Banerjee and Duflo, 2014; Bruhn and Love, 2014; Burgess and Pande, 2005; Fonseca and Matray, 2022; Young, 2017; Cramer, 2021; Barboni et al., 2021; Egger et al., 2022). A complementary approach exploits RCTs to study the implications of access to credit and saving products in developing countries with mixed results (see, for instance, Banerjee et al. (2015) for a review). One explanation for the modest effect of credit access in experimental studies is that they often fail to incorporate the general equilibrium effects. Using combined rich micro-data on firms and labor markets, we are able to directly study the general equilibrium impacts of the bank expansion policy. In the context of India, Young (2017) and Cramer (2021) have examined and validated the bank expansion policy that we use in the paper.[4] Our study focuses on industrial growth and structural transformation, providing new perspectives on the aggregate welfare effects of bank expansion.

Second, our paper contributes to the large literature on structural transformation, especially the impediments to the reallocation of labor from agriculture in developing countries (Gollin et al., 2014; Munshi and Rosenzweig, 2016). There is an ongoing debate about the determinants of structural transformation. Previous research has emphasized the importance of agricultural productivity as the primary "labor push" force (Gollin et al., 2002; Ngai and Pissarides, 2007; Bustos et al., 2016), as well as the importance of human capital growth (Porzio et al., 2022). On the other hand, an older strand of research has suggested the role of the manufacturing sector in the process of structural transformation—the "labor pull" hypothesis. This hypothesis posits that growth in the manufacturing sector could result in higher industrial wages and attract surplus labor from the agricultural sector, thereby driving the structural change (Lewis, 1954; Harris and Todaro, 1970; Bencivenga and Smith, 1997; Gylfason and Zoega, 2006; Alvarez-Cuadrado and Poschke, 2011). Our paper contributes to the literature by providing empirical evidence that manufacturing development acts as a sig-

[4] Cramer (2021) shows that the expansion of banks leads to better health outcomes in India. At the macro level, Young (2017) shows that treated districts had faster economic growth, proxied by nighttime light intensity.

nificant “labor pull” factor, driving the process of structural transformation and potentially reducing labor misallocation.

In addition, our paper contributes to the literature on resource misallocation. A leading explanation of cross-country economic disparity is resource misallocation; however, identifying specific policy tools to reduce misallocation and quantifying their aggregate impacts proves challenging (Hsieh and Klenow, 2009; Bartelsman et al., 2013; Restuccia and Rogerson, 2017; David and Venkateswaran, 2019; Baqaee and Farhi, 2020; Sraer and Thesmar, 2018). In a related paper, Bau and Matray (2023) shows that FDI liberalization in India can reduce misallocation across manufacturing firms. We contribute to this literature by causally examining the effects of banks on the misallocation of resources and talent, both empirically and theoretically, leveraging a rare natural experiment in India's banking system.

The rest of this paper is organized as follows. Section 2 introduces the policy and our identification strategy. Section 3 outlines the data that we use in our analysis. Section 4 presents our findings on industrial growth. In Section 5, we present our main results on structural transformation. Section 6 outlines a quantitative model we can take to the data to study policy counterfactuals. We conclude with Section 7.

2 Bank Expansion Policy and Identification Strategy

2.1 Policy and Institutional Background

The policy reform we analyze in this study was introduced in 2005 by the Reserve Bank of India (RBI) to incentivize banks to open more branches in under-served locations.[5] As per the policy, banks are required to submit an annual branch expansion plan to the RBI, outlining proposed branch openings, closings, and shifts. Thus, by proposing to open more branches in areas that the RBI has designated as “under-banked,” banks can increase their chances of obtaining licenses for their preferred locations. This policy thus effectively “bundled” the entry into under-banked districts with those into more profitable districts.[6] [7]

[5]In India, the banking sector does not permit free entry of banks or bank branches. Banks are required to apply for and acquire licenses from the RBI prior to opening any new branch. Additionally, banks must also request approval to close or relocate branches in most markets.

[6]To make their license-issuance decisions, the RBI also evaluates banks based on other factors, such as the bank's provision of “no-frills” accounts, adherence to priority sector lending obligations, and complaint resolution record. However, these requirements are applied at the bank level, not the individual branch level. For more details on the reform, refer to the 2005 issue of the RBI's *Master Circular on Branch Authorization*.

[7]Banks in India are not allowed to relocate their branches if they leave a market “unbanked”. Therefore, it is not possible for banks to circumvent this policy by opening in under-banked districts and then relocating to bank-rich areas.

The definition of an "under-banked" district is crucial for our identification strategy. According to the rule adopted by the RBI, a district is considered under-banked if the average number of people per bank branch (i.e., the population-to-branch ratio) exceeds the national average for India as follows:[8]

$$\underbrace{\frac{Population_{Dist.}}{\#\ Bank\ Branches_{Dist.}}}_{\text{``}Under-banked\ District\text{''}} > \underbrace{\frac{Population_{National}}{\#\ Bank\ Branches_{National}}}_{\text{``}National\ Average\text{''}}$$

In September 2005, RBI published the list of "under-banked" districts following the above rule, which was then slightly revised in 2006. It is important to note that the RBI did not adjust the list to account for *changes in the ratio*, despite under-banked districts having received more bank branches over time. As a result, the list of "under-banked" districts remained nearly constant throughout our sample period.[9] Using the official RBI document in 2006, we define 375 out of 593 districts as "under-banked" districts, which were spatially dispersed throughout the country.

2.2 Validation of the Policy Cutoff

The arbitrary policy cutoff begets a regression discontinuity (RD) design that compares under-banked districts (treatment group) and control districts with a population-to-branch ratio just above and below the national average. The identification assumption is that districts close to the cutoff are similar in the absence of bank expansion. In this subsection, we show that the policy provides exogenous variation in the presence of bank branches and validate our identification strategy.

To validate this design, we first show that there is no manipulation of the cutoff—so districts do not select into treatment or control groups. The left panel of Figure A.1 presents the histogram plot and non-parametric fitted lines of districts' population-branch ratios (relative to the national average). Visually, there is no sign of bunching on either side of the cutoff, suggesting a limited scope of selection. We formally test it using the McCrary (2008)

[8] According to RBI's *Report of the Group to Review Branch Authorization Policy* published in 2009, the term "national average" refers to a specific statistic provided directly by the RBI. However, the precise methodology that the RBI used to compute this statistic is not explicitly disclosed. To account for this, we independently recalculated this ratio based on its definition, and we validate the accuracy of our calculations in a subsequent section.

[9] While the RBI only published the list of under-banked districts without revealing the detailed district-level population-branch ratios, we reconstruct the ratios using the 2001 population census data and bank branch data from the RBI. After 2010, certain states were made ineligible for "under-banked" status, reducing the number of "under-banked" districts, but no new district was introduced to "under-banked" status. The list of "under-banked" districts was thoroughly updated in 2014 based on the 2011 population census.

density test—the McCrary estimator is -0.065 and the p-value is 0.95. Hence, we cannot reject smoothness around the cutoff.[10]

The right panel of Figure A.1 confirms that the cutoff is indeed meaningful—there is a jump in under-banked status below and above the cutoff. The compliance with the assignment rule is not perfect. Out of 578 districts with bank branch data, seven districts had a status different than predicted by the population-branch ratio. Several reasons might explain the imperfect compliance. First, there may be measurement errors in the branch data.[11] Second, the RBI might have used their discretion to edit the list, potentially with the intent to "help" specific districts, or were captured by political elites. Fortunately, non-compliance does not pose a threat to our analysis, as only 1.7% of observations in our ASI establishment data belong to non-compliant districts. In our main analysis, we exclude non-compliers and use a sharp RD design, assuming perfect compliance. We also demonstrate the robustness of our results by including these observations in a Fuzzy RD design.

We further validate our RD design by showing that other covariates and pre-treatment variables are smooth around the cutoff. Using the 2001 population census data, RBI's bank branch data in 2004, and nighttime light intensity in 2004, Figure A.2a-d visually demonstrate that those district-level characteristics are continuous around the cutoff. We also test for the smoothness of firm-level outcomes in 2004 using the Annual Survey of Industries data. Figure A.2e-h show that firms in treatment and control districts are similar in sales, fixed assets, number of employees, and marginal revenue product of capital prior to policy implementation. Taken together, these results suggest that districts are properly randomized around the cutoff, lending support to the causal interpretation of our RD results.

[10]The lack of manipulation is not surprising as the banking system in India is tightly regulated. By the *1949 Banking Regulation Act*, banks should submit detailed annual expansion proposals and cannot, without the prior approval of the RBI, open new branches or change the location of existing branches. Recall that the population-branch ratio has two components: 1) the district population in 2001, which was fixed when the expansion policy was announced; and 2) the number of bank branches in the district. To game the policy, banks had to collectively gain RBI's approval to open or close branches in certain districts at least one year before the policy. It is highly unlikely, as the policy was announced in 2005 without prior notice.

[11]The total number of districts in 2001 was 593. For districts that were split from a 2001 district, we recoded them to the original district. Some districts were excluded because they were formed after 2001 by merging several existing districts, making it impossible to map them to previous districts. After a careful review of RBI documents and historical texts, we further excluded three districts due to suspected coding errors or manipulation. Ujjain, a historically wealthy city not initially included in the list of under-banked districts in 2005, was added in 2006. We suspect this addition might have been politically motivated, and hence, we excluded Ujjain. Badgam district was also removed from our sample. There was a transfer of lead bank responsibility in respect of Anantnag, Budgam, Pulwama and Srinagar districts to the Jammu & Kashmir Bank Ltd. up to March 2005. These bank branches were recorded as closed in the RBI bank branch data, reversing the treatment status of Badgam. Lastly, we dropped Varanasi due to the 2002 merger of the private sector Banaras State Bank with the nationalized Bank of Baroda. This merger led to some bank branches in Varanasi being coded as closed, altering their treatment status.

2.3 Identification Strategy: Difference-in-Discontinuity

To capitalize on the temporal dimension of our bank branch and firm datasets, as well as the sudden shift in policy, we implement the Difference-in-Discontinuity (Diff-in-Disc) design, in line with the methodology outlined in Grembi et al. (2016). Intuitively, this approach initially compares treatment and control districts with a population-per-branch ratio near the policy's threshold, as in the standard Regression Discontinuity (RD) design. Subsequently, it evaluates the differences in this discontinuity before and after policy implementation, mirroring the Diff-in-Diff design. By employing this Diff-in-Disc approach, we can transparently track and present the dynamic impacts of the policy over time.

To operationalize, we estimate the following empirical equation:

$$\begin{aligned} y_{(i)dt} =& \beta_1 UnderBank_d + \beta_2 Ratio_d + \beta_3 UnderBank_d * Ratio_d \\ & + Post_{2006} * (\delta_1 UnderBank_d + \delta_2 Ratio_d + \delta_3 UnderBank_d * Ratio_d) \\ & + \sigma_d + \sigma_t + X_{(i)dt} + \varepsilon_{(i)td} \\ s.t. & -h < Ratio_d < h \end{aligned}$$

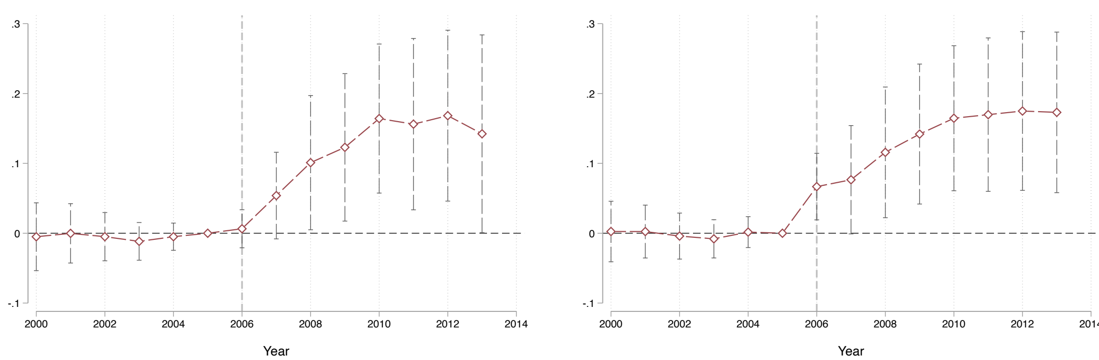
where $y_{(i)dt}$ includes outcomes of firm i in district d in year t in our firm-level regression, and district-level outcomes, such as the number of branches and total manufacturing employment, in district regression. The first three terms on the right-hand side comprise the standard RD design with the running variable $Ratio_d$—the population-branch-ratio of district d. By fitting a first-order polynomial of the running variable on both sides of the cutoff, β_1 captures the effect of being in treated districts on the outcome variable $y_{(i)dt}$. We then interact them with a post-policy dummy variable in the second row, allowing for changes in the discontinuity due to the policy. δ_1 thus captures the treatment effect of the bank expansion policy in under-banked districts.

To address potential issues that districts with varying population-to-branch ratios might exhibit different trends in the absence of the policy, we control for components of the ratio in the vector $X_{(i)dt}$. This includes the pre-policy district population and the number of branches, both interacted with a linear time trend. For our firm-level regression, $X_{(i)dt}$ also includes a set of firm- and industry-level covariates, including firm ownership fixed effects, an urban dummy variable, as well as 2-digit industry-by-year and state-by-year fixed effects in some specifications. Additionally, σ_t and σ_d represent year and district fixed effects, respectively, to control for temporal and geographical variations. Importantly, we always control for district fixed effects in our regression, so any time-invariant district-specific shocks and noise cannot bias our estimates. Standard errors are clustered at the district level to account for potential correlation within districts. The parameter h is the MSE-optimal bandwidth and only districts within this bandwidth are included in the sample. We conduct several robustness checks to confirm the consistency of our results across different bandwidths.

It is important to emphasize that the underlying assumption of our identification strategy is the existence of a *local* parallel trend among firms in districts marginally below and above the national average. This premise is akin to what is typically assumed in the standard Difference-in-Difference (DiD) design. Notably, this assumption is less stringent than those in the standard Regression Discontinuity (RD) design, which requires smoothness across all covariates around the cutoff. While we have tested this smoothness requirement for several key covariates, as illustrated in Figure A.2, our approach requires only that any discontinuities at the cutoff, irrespective of their origin, remain consistent in the absence of the policy. We will further validate this parallel trend assumption with event study plots in the subsequent analysis.

2.4 Effectiveness of the Policy

The reform proved to be highly effective, creating strong incentives for banks to establish a presence in districts previously deemed unprofitable and thus under-banked. Utilizing the standard Regression Discontinuity (RD) design, both Cramer (2021) and Young (2017) have demonstrated a marked response by banks to this policy. We replicate these findings with our preferred Difference-in-Discontinuity (Diff-in-Disc) specification and illustrate the treatment effects on the number of bank branches and branch licenses in the event study plots presented in Figure 1.



(a) log(Branches) (b) log(Licenses)

Figure 1: Event Study Graphs for the Treatment Effects on Bank Branches and Licenses

On average, treatment districts received 21% more branch licenses and 19% more branches than control districts by 2010, which corresponds to an increase to 8.31 branches per 100,000 people, compared to the control mean of 6.99 branches. Furthermore, treatment districts also saw a large increase in deposits and credit after the policy.[12]

3 Data

Our primary source of data is the establishment-level data from the Annual Survey of Industries (ASI) from 1998 to 2013.[13] The ASI provides a representative sample of all registered manufacturing establishments in India, with large establishments covered yearly and smaller establishments surveyed on a sampling basis. The publicly available ASI includes unique plant identifiers that are consistent across the years starting from 1998. However, it lacks district information, which is critical to our analysis. To address this, following Martin et al. (2017), we match the panel version of ASI with an older cross-sectional version, which contains district identifiers until 2009, based on time-invariant factors and open/close variables.[14] The ASI includes comprehensive plant-level information on revenues, labor costs, stock of fixed assets, and materials, among others, which are essential for constructing our key firm outcome variables. We perform substantial data cleaning and deflate all nominal outcome variables to constant 2004-2005 Rupee following Allcott et al. (2016).

Our labor market data come from the Socioeconomic High-resolution Rural-Urban Geographic Dataset (SHRUG),[15] which integrates data from multiple rounds of the population and economic census. In particular, we use the 2001 and 2011 village-level Population Census data from the Primary Census Abstract and Village Directory tables. This data provides information on village infrastructure, demographics, employment, occupation, and population, which are used to construct variables related to labor supply in agricultural and manufacturing & service sectors. It also provides a basis for connecting all other datasets at the village level. We complement this with crop suitability data from the FAO Global Agro-Ecological Zones (GAEZ). This dataset assesses crop suitability and production potential based on plant characteristics, climate, and soil quality, aggregated to the village level by SHRUG.

[12] The increase in deposit and credit is especially pronounced for private sector banks, as depicted in Figure A.3. One concern is that private banks might simply "steal" market share from existing public sector banks, which could lead to minimal aggregate effects in local markets. However, our results in Figure A.4 suggest that although the stealing effect might be possible, its effects tend to be small and insignificant.

[13] The reporting period for the ASI is the Indian fiscal year, which begins on April 1 and ends on March 31. Throughout the paper, when we refer to a survey year, we use the calendar year in which the fiscal year commences. All financial amounts are expressed in 2004 Rupees.

[14] Since the ASI does not provide district identifiers after 2009, we use the panel structure of our data to infer the district information for a subset of firms (approximately 2/3) that appear in the data before 2010.

[15] Asher and Novosad (2020) provides details of the data construction, accessible at https://www.devdatalab.org/shrug.

Data pertaining to the banking sector and the implementation of the policy is obtained from the Reserve Bank of India. The list of under-banked districts is digitized from the report of RBI in 2006. While the exact district-level population-branch ratios are not included in the report, we are able to reconstruct them using the 2001 population census data and bank branch data from the Master Office File published by the RBI. We also obtain district-by-bank-group level credit and deposit data from 2003 to 2016 from the RBI. Since the list of under-banked districts is based on the 2001 population census districts (and remains unchanged until 2014, when the list was updated according to the 2011 population census), we build a crosswalk to map all data to the 2001 population census districts.[16]

4 Did Bank Expansion Lead to Industrial Growth?

In this section, we examine the effects of bank expansion on manufacturing firms. Starting with a visual exploration of cross-sectional and time-series plots using raw data from the ASI, we identify key trends and variations in firm size across under-banked and control districts. Subsequently, we detail our identification strategy, employing a Difference-in-Discontinuity design that leverages the policy cutoff and its implementation timeline to isolate the causal effects of bank expansion. Finally, we present our empirical findings at the firm and district levels, providing insights into the micro and macro impacts of bank expansion on firm dynamics, labor allocation, and industrial growth.

4.1 Who Are the Under-Banked?

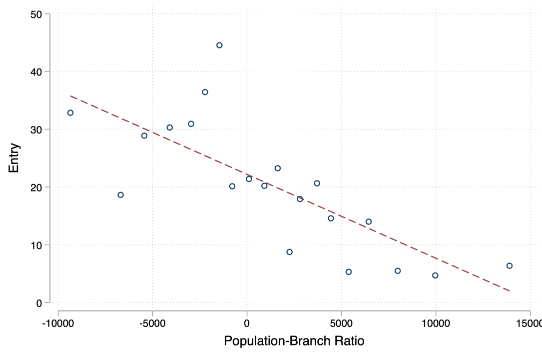
We start by examining the correlation between districts' population-per-branch ratio (the metric that the RBI used to define under-banked districts) and local industrial development before 2005. Figure 2 illustrates this ratio against several pre-reform district-level measures of industrial development and productivity, including the number of new entrants, logged number of plants, manufacturing wages, and Total Factor Productivity (TFP).[17] The data reveal that under-banked districts tend to have fewer new entrants and a smaller number of plants. This lack of manufacturing activity significantly impacts the local labor market, leading to lower wages in these districts. Interestingly, under-banked districts exhibit higher average TFP. This counterintuitive finding could be explained by the presence of higher entry

[16] The district borders in India are very volatile. There were 593 districts in 2001 and 640 districts in 2011.

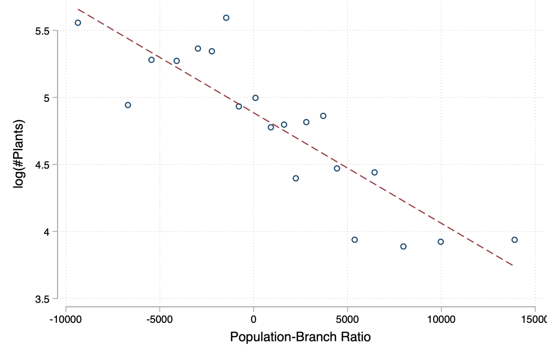
[17] The number of entrants is not logged due to the presence of ero in many district-year observations. District-level TFP is measured by the average Solow residuals across all firms in the district, using the sampling weights. Year fixed effects are projected out.

or fixed operating costs in under-banked areas, which might deter less productive firms from entering the market, leaving only the more productive ones.[18]

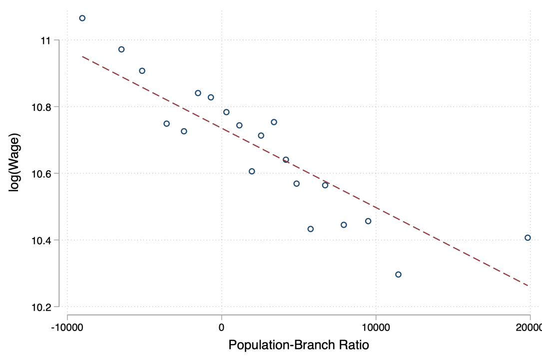
It is important to recognize that these correlations should not be interpreted as causal relationships. The lack of bank branches in these districts might contribute to the under-development of the manufacturing sector. However, it is also plausible that other underlying distortions, such as labor market frictions or inefficient industrial policies, are at play. These distortions might reduce the demand for credit in these districts, making them appear under-banked. To differentiate these two explanations, we move to the time trends of firm size distribution in under-banked vs. control districts.



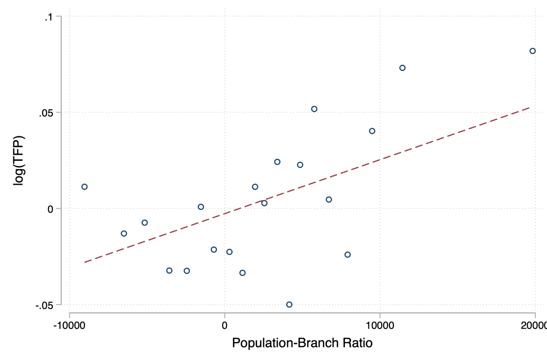
(a) Entry



(b) log(Num. of Plants)



(c) log(Wage)



(d) log(TFP)

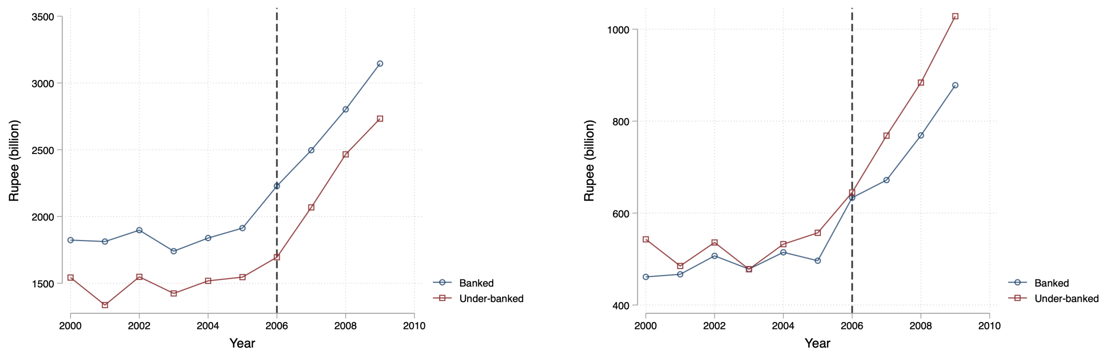
Figure 2: District Population-Branch Ratio and Industrial Development

4.2 Firms in Under-banked and Banked Districts: Time Trends

Total Capital. Figure 3 displays the time trends of total capital in all under-banked and banked districts (left panel), and focusing exclusively on districts around the policy cutoff

[18]This explanation aligns with the heterogeneous firm model proposed by Melitz (2003), where a fixed export cost results in positive selection into the export market.

(right panel). Total capital is constructed as the weighted sum of firm-level fixed assets[19] in the treatment and control groups, using the survey sampling weights. Capital is deflated to constant 2004-2005 Rupee values using the Gross Capital Formation data from the RBI.



(a) All distircts

(b) Districts around the policy cutoff

Figure 3: Trends of Total Fixed Assets in Treatment and Control Districts

The left panel reveals that under-banked districts had approximately 25% less capital than banked districts. This gap remained stable despite rapid industrial growth from 2000 to 2005. Under-banked districts began to catch up following the policy implementation as new bank branches entered and more credit was issued. This catch-up is more evident when focusing on districts around the policy cutoff, as displayed in the right panel. Under-banked and banked districts had remarkably similar levels of capital prior to the policy, suggesting that the treatment status assignment was as good as random around the policy cutoff. This further validates our RD design. However, divergence began as those narrowly under-banked districts received more bank branches.

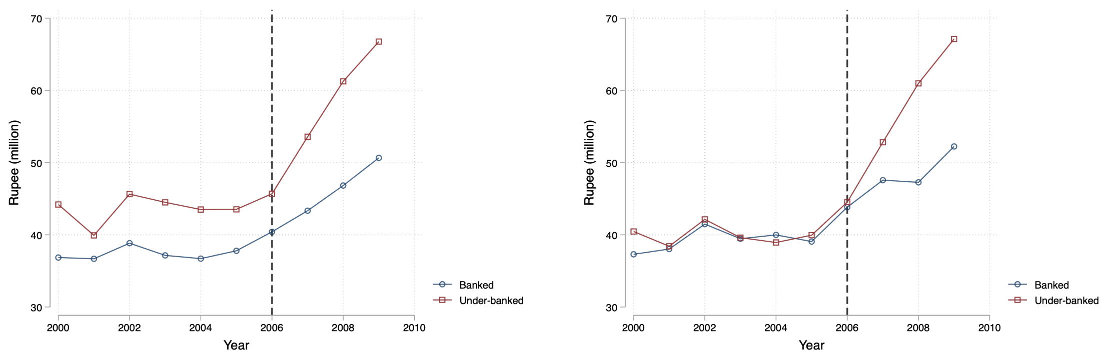
Average Firm Size. These patterns interestingly invert when examining the capital of the average firm, as shown in Figure 4. The left panel indicates that before the policy, firms in under-banked districts were approximately 20% *larger* than firms in banked districts in terms of fixed assets. This pattern is consistent with our cross-sectional evidence that under-banked districts tend to have higher TFP due to the selection of more productive firms.

There are two potential explanations for the positive selection. Firstly, the lack of access to finance might *cause* high costs of operation, such as a fixed cost of using credit, as posited by Ji et al. (2023). Consequently, only the more productive and talented entrepreneurs would bear the costs and enter the market. In this case, we would expect that opening more bank

[19]Fixed assets include tangible assets such as plants, land, and machinery owned by firms, but exclude mining rights and other intangible assets.

branches in under-banked districts could lower the operating costs, promoting the entry of smaller (and possibly less productive) firms (Bazzi et al., 2021), thereby reducing the average firm size.

Alternatively, the larger average firm size and the scarcity of bank branches could both be the *result* of inefficient industrial policies or other structural barriers. Suppose firms in under-banked districts are less profitable due to other inefficient policies or barriers, then only a few talented entrepreneurs enter and stay in the market. The aggregate demand for credit is suppressed due to fewer operating firms. This muted credit demand could then discourage the entry of banks into these districts, rendering them "under-banked". In this scenario, simply opening more bank branches might not promote the entry of potential entrepreneurs.



(a) All distircts

(b) Districts around the policy cutoff

Figure 4: Trends of Firm Average Fixed Assets in Treatment and Control Districts

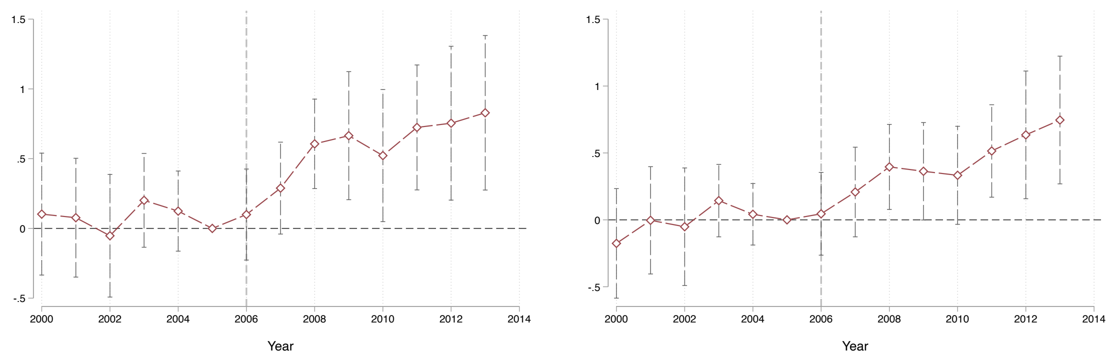
Proceeding to the post-2006 time-series, both panels indicate that firms in under-banked districts grew even larger following the policy, aligning more with the second hypothesis. This outcome yields two critical insights. Firstly, despite firms in under-banked districts being positively sorted and consequently larger than their counterparts in banked districts, they remain too small possibly due to financial constraints. Secondly, bank expansion seems to alleviate these constraints more substantially along the intensive margin than the extensive one. As a result, the reform appears to have primarily driven the growth of incumbent firms rather than fostering the entry of potential entrants.

This discussion underscores a key tension in the debate on financial inclusion: is "under-banked" a symptom or a cause? In one scenario, under-banked districts might face high *bank* entry or operational costs due to factors like stringent financial regulations. Financial inclusion programs, if properly implemented, could alleviate these frictions, thereby promoting bank entry and spurring economic growth. However, if both the "under-banked" status and lack of economic development are driven by other distortions, such as labor market frictions or inefficient industrial policies, merely increasing the number of banks may not address the underlying issues. With this question in mind, we now proceed to examine the causal effect of bank expansion on manufacturing development.

4.3 Effects of Bank Expansion on Firms

In this subsection, we examine the treatment effects of bank expansion on average firm size, entry, and exit. In addition, we will also explore the aggregate effects at the district level. To capture the dynamic effects of the policy, we employ the Difference-in-Discontinuity design as outlined in Section 2.3.

Event study graphs. Figure 5 displays the event study graphs for capital and sales revenue. These graphs report the yearly treatment effects of being situated in (near cutoff) treated districts relative to the controls, using the same controls and bandwidth as our baseline equation. The nonexistence of a discernible effect before the reform provides visual evidence of the "parallel trends" assumption, thereby validating our identification strategy.[20] Figure A.7 reports similar patterns and corroborates the absence of pre-trends for wage bills and total employment.



(a) log(Fixed Assets)

(b) log(Sales Revenue)

Figure 5: Event Study Graphs for the Treatment Effects on Capital and Sales

Following the bank expansion, firms in treated districts grow larger by using more capital and labor, and generating more revenues. These effects are both economically significant

[20] A potential concern could be the migration of plants from control to treated districts, in anticipation of improved financial access. However, such a movement is implausible given India's stringent industrial regulation scheme. Evidence from our plant panel data from 1998 to 2013 indicates that less than 5% of all plants have ever relocated their districts. Note that this figure could be inflated due to possible measurement or coding errors in the district information. Our baseline analysis uses a plant's modal district as the time-invariant district. Our results remain robust when plants that have relocated are excluded.

and unfold progressively over time, in line with the idea that changes in the allocation of resources are typically slow-moving. Meanwhile, it is worth noting that the adjustments in labor appear to lag even further, likely due to other constraints in the labor markets, as documented by Bertrand et al. (2021).

Baseline estimates. Table 1 presents the estimated treatment effects of the bank expansion policy on firm-level sales revenue, fixed assets, and wage bills using our baseline estimating equation. For the average firm, capital increases by 37% (column 2), indicating that the policy exerts large positive effects on capital investments.[21] The higher capital investment does not crowd out labor, as wage bills and employment increase by approximately the same amount (columns 3-4), suggesting strong complementarity between capital and labor. Post-reform, firms in treated districts expand their sales revenue by approximately 30%.

Table 1: Treatment Effect of Bank Expansion on Firms

Dependent Variable	(1) Revenues	(2) Capital	(3) Wage Bill	(4) Employment
Treated * Post	0.291**	0.371**	0.357***	0.306***
	(0.142)	(0.166)	(0.133)	(0.113)
Observations	135,673	135,673	135,673	135,673
R-squared	0.173	0.212	0.199	0.147
District FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Firm Controls	Yes	Yes	Yes	Yes
District Trends	Yes	Yes	Yes	Yes

Notes: All outcome variables are in logs. Firm Controls include firm ownership fixed effects and a dummy variable of being in urban areas. District Trends include district population in 2001 and number of bank branches in 1997, interacted with a linear time trend. Standard errors are clustered at the district level. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Robustness. Table A.1 shows that the estimates are robust under the most parsimonious specification, which includes only district and year fixed effects. As we demonstrate in Table A.2, our results are robust to the inclusion of 2-digit industry-by-year fixed effects. By comparing firms in the same 2-digit industry in the same year, this specification accounts for any unobserved, time-varying, sector-level shocks, such as aggregate trade shocks and changes

[21] As shown in A.3, the increase in fixed assets is mostly driven by investments in buildings, plants, and machinery, affirming that firms use newly acquired credit to improve production capacity.

in the priority sector lending policies at the 2–digit industry level.[22] In addition, to account for the possibility that some Indian states are more exposed to the reform and may have adjusted their state-level banking regulations or been affected by other concurrent shocks, we flexibly control for any state-level time-varying unobserved shocks. Table A.2 shows that our results are robust to the inclusion of state-by-year fixed effects. Moreover, Figure A.8 and A.9 demonstrate that our results hold robustness to dropping individual states and 2-digit industries one by another. Figure A.10 shows that using different bandwidths does not change our estimates qualitatively.

Outstanding loans. One natural question is to what extent the growth of firms documented above is attributable to bank expansion alleviating credit market frictions. An alternative hypothesis is that bank expansion might stimulate local economic activity and consumption, allowing firms to generate higher sales revenue, which they can reinvest. While we ideally want to examine the treatment effect on loans that firms obtained directly from banks, the ASI data only provides information on total outstanding loans without specifying their sources. Nevertheless, we present the effect on total outstanding loans in Table A.4 (column 1-3) and the event study graph in Figure A.7 (panel (c)). Following the reform, outstanding loans experienced a 41% increase (using our baseline equation in column 1) for the average firm in treated districts.[23] The substantial increase in total outstanding loans suggests a real expansion in firm-level credit rather than a reorganization of existing liabilities—using bank loans to replace other higher-interest debts. Notably, this point estimate aligns with the estimate for fixed assets, implying that the alleviation of borrowing constraints can account for the entirety of the observed firm growth.

Product Scope. Another dimension to consider is that as households become richer, they may diversify their consumption portfolio by purchasing more varieties of products (Li, 2021). This increase in the demand for variety could encourage firms to expand their production

[22] All banks (public and private) are required to lend at least 40% of their net credit to the "priority sector", which includes agriculture, agricultural processing, transport industry, and small scale industry (SSI). If banks fail to satisfy the priority sector target, they are required to lend money to specific government agencies at very low interest rates (Banerjee and Duflo, 2014). The definition of the priority sector has expanded over time. There are 53 distinct 2-digit industries in our ASI data.

[23] In our sample, more than 85% of firms reported having positive outstanding loans even before the reform. Considering that most firms had already accessed some form of loan, we did not observe a significant effect of bank expansion on the binary variable representing loan usage, as shown in Table A.4 (column 4-6). However, it's crucial to highlight that this finding does not rule out the possibility that bank expansion could reduce the fixed costs of accessing credit. If credit usage is a prerequisite for market entry, we would expect all existing firms to borrow. In this scenario, a reduction in borrowing costs would not change the proportion of firms using loans; instead, it could lower entry barriers and alter the composition of the firm pool.

by introducing new products. This discussion ties into a broader debate in growth theory (Garcia-Macia et al., 2019): do firms achieve growth through the creation of new products, the process of creative destruction, or by improving their existing products? If relaxing financial constraints could lead to the improvement of existing products or the introduction of new products, the bank expansion reform could yield substantial dynamic gains by stimulating innovation.

We can directly test the effects on product composition by leveraging a unique feature of the ASI, which reports both total product sales and total quantity sold at the firm-product level, as mandated by the 1956 Companies Act.[24] With this information, we compute the establishment-level total number of products, a price index constructed as the weighted average of product prices, and indicators of product addition and deletion using the firm-product panel.

Table 2: Treatment Effect of Bank Expansion on Product Portfolio

	(1)	(2)	(3)	(4)
Dependent Variable	log(Product)	log(Price)	Add Product	Del. Product
Treated*Post	0.029	-0.048	-0.004	-0.003
	(0.035)	(0.188)	(0.023)	(0.018)
Observations	123,972	123,871	86,286	86,286
R-squared	0.093	0.104	0.025	0.021
District FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Firm Controls	Yes	Yes	Yes	Yes
District Trends	Yes	Yes	Yes	Yes

Notes: log(Product) is the log number of products, and log(Price) is the firm-level average product price, weighted by product sales revenue. Add Product is a dummy variable equal to 1 if the firm has more products than the previous year. Del. Product equals 1 if the firm has fewer products than the previous year. Firm Controls include firm ownership fixed effects and a dummy variable of being in urban areas. District Trends include district population in 2001 and the number of bank branches in 1997, interacted with a linear time trend. Standard errors are clustered at the district level. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 2 shows that the reform has a negligible and statistically insignificant impact on the number of products (column 1), contradicting the demand for variety hypothesis. Column 2 reveals that bank expansion exerts a slightly negative, yet not significant, effect on the price index, hinting at limited improvement in product quality. The muted effect on price also alleviates another concern that our results might be driven by an increase in demand in treated districts, which should push up output prices. Furthermore, columns 3-4 indicate that

[24]The Act requires Indian firms to disclose product-level information on capacities, production, and sales in their annual reports. The product is defined based on the 5-digit product codes and comprises 11,880 distinct products.

firms in treated districts are not more likely to add or delete products post-reform compared to their counterparts in control districts. Thus, the observed firm expansion is unlikely to be driven by firms creating new products or stealing from others. Instead, the finding suggests firms primarily grow by increasing the production of existing products, hinting at limited dynamic gains.

Firm Entry and Exit. While we have demonstrated that the average firm in treated districts grew larger following the reform, it remains unclear whether it is driven by the growth of firms or changes in the *composition* of firms. We construct firm entry and exit indicators using our panel data, following the methodology of Harrison et al. (2015).[25] Both entry and exit rates in our sample are at 6.7%,[26] indicating that the economy is in a steady state. Table 3 presents the treatment effects on firm entry in Panel A and exit in Panel B. The first column reveals that the reform has positive but statistically insignificant effects on both entry and exit rates. This suggests that our observed firm expansion results are primarily driven by the growth of incumbent firms, rather than changes in the composition of firms.

However, a closer examination reveals substantial variation across firm sizes. Column 2 presents the heterogeneous treatment effects, in which our "Treated*Post" term is interacted with a "Big" dummy variable, assigned as 1 if a firm's average employment surpasses the national average.[27] The main "Treated*Post" term suggests an increase in the exit rate of smaller firms by 4.7 percentage points (approximately 70% of the sample mean). This aligns with our theoretical framework, as the reform raises the exit rate among smaller firms due to the elevation of the unconstrained productivity threshold. Interestingly, this main effect is almost entirely offset by the interaction term, indicating that larger firms' exit rates are unaffected by the reform. Regarding firm entry, the reform also exerts a slightly larger, albeit still insignificant, effect on the entry of larger firms, consistent with our model. To further validate our findings, we present additional results in columns 3-5 using the baseline Difference-in-Discontinuity specification, disaggregated by firm sizes. These results confirm that the observed higher exit rate is mainly driven by very small firms (with fewer than 20

[25] An entry is defined as a firm appearing in the data for the first time within three years of the initial production year. An exit is when a firm is officially declared "closed" in the ASI and remains so.

[26] The figures are consistent with the exit rate imputed from plant age cohorts by Hsieh and Klenow (2014), also using the ASI data.

[27] We use the average establishment size over the years as a proxy for firm-level productivity. Given that firms usually enter the market at a smaller size, and considering that some firms may already be on a downward trajectory before their exit, we opt not to utilize the establishment size at the points of entry and exit. This "Big" dummy is interacted with all single and cross-terms in our baseline Difference-in-Difference specification.

Table 3: Treatment Effect of Bank Expansion on Firm Entry and Exit

	Panel A: Entry				
	(1)	(2)	(3)	(4)	(5)
	Full Sample	Full Sample	$L \in (0, 20)$	$L \in [20, 100)$	$L > 100$
Treated*Post	0.009	0.001	0.002	0.024	-0.001
	(0.010)	(0.016)	(0.018)	(0.017)	(0.010)
Treated*Post*Big		0.009			
		(0.021)			
Observations	166,094	166,094	51,623	52,553	61,917
R-squared	0.047	0.054	0.062	0.054	0.032
Mean of dependent variable	0.0674	0.0674	0.0918	0.0867	0.0306

	Panel B: Exit				
	(1)	(2)	(3)	(4)	(5)
	Full Sample	Full Sample	$L \in (0, 20)$	$L \in [20, 100)$	$L > 100$
Treated*Post	0.013	0.047**	0.069**	0.033	0.003
	(0.012)	(0.022)	(0.027)	(0.025)	(0.008)
Treated*Post*Big		-0.038*			
		(0.020)			
Observations	166,094	166,094	51,623	52,553	61,917
R-squared	0.027	0.045	0.053	0.040	0.021
Mean of dependent variable	0.0674	0.0674	0.101	0.0768	0.0314

Notes: Entry equals 1 in the first year an establishment appears in the data within three years of the initial production year. Exit equals 1 if an establishment is officially declared "closed" in the ASI and remains closed thereafter. All regressions include district and year fixed effects, firm controls, and district trends. Big is a dummy variable equal to 1 if the establishment's average employment number is greater than the national average. We interact this Big dummy with all the single and cross-term in our baseline Diff-in-Desc specification. Firm controls include firm ownership fixed effects and a dummy variable of being in urban areas. District trends include district population in 2001 and the number of bank branches in 1997, interacted with a linear time trend. Standard errors are clustered at the district level. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

employees), while the effects on larger firms are small and insignificant.

TFP. We conclude our firm-level analysis by noting that the reform had minimal effect on firm productivity. We measure firm-level productivity (TFP) using two methods: (1) calculating Solow residuals, and (2) estimating the revenue production function, following the approach outlined by Levinsohn and Petrin (2003). Table A.5 presents the effects on TFP using both measures. Interestingly, our results indicate that the reform has precisely estimated null effects on both measures of firm productivity. This finding aligns with our static model that abstracts from innovation and previous empirical results, indicating that incumbent firms grow by employing more resources to increase the production of existing products under the same technology. In comparison, changes in firm composition and innovation (including both product and process innovation) play a relatively insignificant role.

4.4 District-level Aggregated Effect

Our analyses above have shown that the average firm expands due to bank expansion. In this subsection, we present district-level results to explore the aggregate effect of the reform on local markets, echoing our graphic evidence reported in Figure 3. Our district-level variables are constructed by aggregating the establishment-level variables, inflated by their sampling weights. We use the same bandwidth as in our plant-level analysis with the same estimating equation to ensure that we are comparing the same set of treatment and control districts.[28]

To further validate our empirical strategy, Figure A.11 presents the event study plots for district-level aggregate revenues, capital, wages, and employment. Consistent with the time-series plots of total capital in Figure 3, treated districts did not experience faster industrial growth prior to the policy implementation, providing visual evidence that pre-trends cannot bias our estimates.

Table 4 presents our Difference-in-Discontinuity results at the district level. Importantly, these results align closely with the firm-level estimates reported in Table 1. The coefficients suggest a significant industrial growth in districts that were narrowly under-banked prior to the reform. Post-reform, these treated districts saw a 39% surge in manufacturing output, facilitated by a substantial increase in fixed assets (39%) and manufacturing labor—demonstrated by a 34% rise in wage bills (column 3) and a 31% growth in employment (column 4).

The remarkable consistency between the district-level aggregate results and the firm-level findings suggests that changes in firm composition have a negligible impact on the observed trends. This is further supported by the insignificant effects on firm entry and exit documented in Table 3.

[28]We replace our firm-level controls, including firm ownership fixed effects and an urban dummy variable, with the district share of private firms and the share in urban areas. Some districts have very few firms in some years and generate a significant amount of noise; hence, we drop district-years with less than 20 firms. Dropping these observations does not change our estimates. In addition, we control for the number of plants in a district-year to account for the noise, which, again, does not change our estimates.

Table 4: Treatment Effect of Bank Expansion on District Aggregate

	(1)	(2)	(3)	(4)
Dependent Variable	Revenue	Capital	Wage Bill	Employment
Treated * Post	0.385**	0.387***	0.340*	0.308**
	(0.183)	(0.136)	(0.172)	(0.126)
Observations	1,838	1,838	1,838	1,838
R-squared	0.960	0.935	0.972	0.969
District FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
District Trends	Yes	Yes	Yes	Yes

Notes: All outcome variables are constructed as the weighted sum of firm-level variables using the sampling weights and then transformed in logs. District Trends include district population in 2001 and the number of bank branches in 1997, interacted with a linear time trend. Additional controls include the share of private firms, the share in urban areas, and the number of firms in a district-year. Regressions are weighted by district-level capital size during the pre-reform period. District-year observations with less than 10 plants are Dropped. Standard errors are clustered at the district level. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

5 Effects on Structural Transformation

So far, we've shown that manufacturing firms experienced significant expansion in treated districts following the reform. In this section, we further explore the implications of this growth on local labor markets, particularly focusing on the extent to which banking sector reforms shifted labor from agricultural to manufacturing sectors. On one hand, the growth of manufacturing firms, acting as a pull-side factor, creates new industrial job opportunities and elevates wages in the manufacturing sector. As a consequence, the reform might "trickle-down" to poor households and draw workers away from agriculture.[29] On the other hand, improved access to finance could also stimulate agricultural investments, potentially increasing the marginal productivity of agricultural labor.[30] Hence, the net effect on inter-sectoral labor reallocation is *ex-ante* unclear.

Furthermore, existing literature has highlighted substantial inter-sectoral labor mobility

[29]In a similar vein, Barboni et al. (2021) investigates the effects of bank expansion also in the context of India, leveraging experimental evidence. Their study reveals that while bank expansion facilitated greater access to loans for impoverished rural households, the recipients primarily used these loans for consumption purposes rather than for investment. Despite this consumption-focused utilization of loans, these households experienced an increase in rural income. This outcome is potentially attributed to the "trickling-down" effect: the heightened economic activity engendered by the expansion of banking services indirectly catalyzes an increase in local wages.

[30]Agricultural investments and new technology can potentially lower the marginal productivity of labor, as exemplified by the case of Brazil's GE soybean seeds (Bustos et al., 2016). However, Madhok et al. (2022) demonstrates that, in India, most agricultural investments serve to augment labor rather than replace it.

costs in India, potentially confining labor to agriculture even in the face of low productivity. Munshi and Rosenzweig (2016) highlights that the wage gap between urban and rural areas in India is substantial, with urban wages being over 47% higher than rural wages even for unskilled workers. This wage gap remains constant over time and is significantly larger than in other developing countries, such as China and Indonesia. Surprisingly, despite the potential for higher wages in urban areas, rural workers in India do not capitalize on these arbitrage opportunities, as evidenced by the country's low migration rate.[31] Several factors may impede labor mobility and lead to the misallocation of labor in India, including limited access to transportation infrastructure such as roads (Asher and Novosad, 2020) and the informal insurance networks deeply rooted in the traditional caste system (Munshi and Rosenzweig, 2016). Consequently, it remains an empirical question whether industrial growth can effectively release workers trapped in the agricultural sector, particularly those with lower agricultural productivity.

In this section, we illustrate that bank expansion, by fueling industrial growth and creating more non-agricultural job opportunities in treated districts, can alleviate labor misallocation. Farmers in villages with lower agricultural suitability strongly respond to the reform and transition away from agriculture. To quantify this effect, we utilize the village-level population census data from 2001 and 2011,[32] matched with crop suitability data from the FAO Global Agro-Ecological Zones (GAEZ)[33] as facilitated by Shrug (Asher and Novosad, 2020). Villages are sorted into quartiles based on their crop suitability *within* a sub-district, where they likely encounter similar industrial labor demands.

We then evaluate the treatment effect of bank expansion on the number of farmers in each quartile separately. More specifically, we estimate the following empirical equation using the local linear RD approach with MSE-optimal bandwidth, following the methodology proposed by Calonico et al. (2014).

$$Y_{ids} = \beta_1 UnderBank_{ds} + \beta_2 Ratio_{ds} + \beta_3 UnderBank_{ds} * Ratio_{ds} + X_{ids} + \sigma_s + \varepsilon_{ids}$$

[31] Drawing on detailed district-wise migration flow data from the 2001 population census, Kone et al. (2018) underscores that the 5-year inter-district migration rate in India is remarkably low, at a mere 2.8%. In stark contrast, the 5-year inter-prefecture migration rate in China stands significantly higher at 10%. In light of these facts, spatial sorting of workers according to skills a la Young (2014) is insufficient to explain the substantial rural-urban wage gap in India.

[32] Villages in India are typically small, with approximately 500,000 in total, each averaging around 1,000 inhabitants. Over 70% of all main workers are engaged in the agricultural sector.

[33] The GAEZ crop suitability data has been extensively applied in the literature, as reviewed by Donaldson and Storeygard (2016), among others. GAEZ utilizes agronomic models to predict potential crop yields based on location characteristics (including topography and soil type) under various levels of input usage.

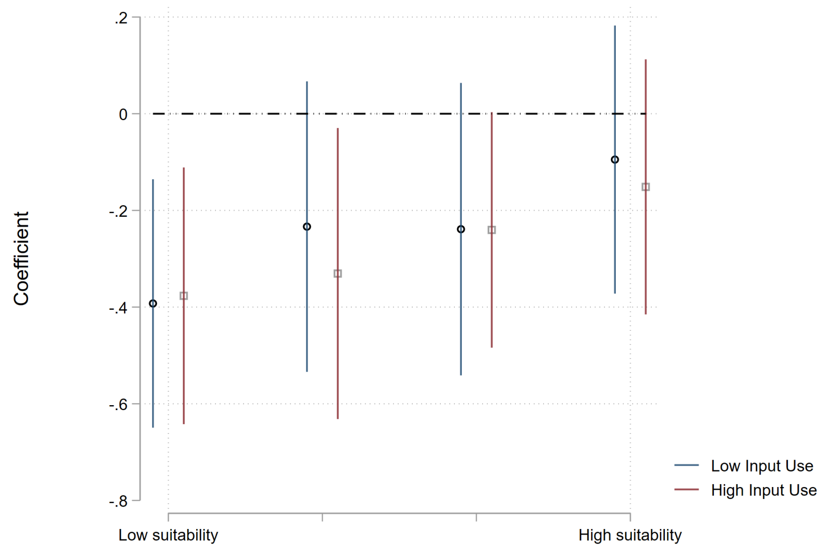


Figure 6: Treatment Effect of Bank Expansion on log(Farmer) by Crop Suitability

$$s.t. -h < Ratio_{id} < h$$

Here i denotes village, d denotes district, and s denotes state. Y_{ids} includes our village-level outcome variables in 2011. $Ratio_{ds}$ is district d's population-branch ratio (relative to national average) and $UnderBank_{ds} = 1$ if district d that has a $Ratio_{ds} > 0$. We control for the ratio's components, village's land area, and population in 2001 in X_{ids}, and state fixed effect in σ_s. We also control for the baseline measures (using the 2001 population census) of the respective variable of interest as recommended by Lee and Lemieux (2010). h is the estimated MSE-optimal bandwidth. Standard errors are clustered at the district level.

Figure 6 illustrates the treatment effect on the log number of farmers by cereal crop suitability quartiles. The number of farmers decreases by roughly 40% in villages with the lowest crop suitability, as demonstrated on the left side of the plot. As we move to villages with higher crop suitability, the treatment effect gradually diminishes in magnitude and becomes statistically indistinguishable from zero. These findings align with the hypothesis that bank expansion reduces labor misallocation by encouraging labor mobility away from the agricultural sector in villages with lower agricultural productivity.

The calculation of crop suitability requires an assumption about input usage.[34] While the farming system in India typically aligns more closely with the low-input case (represented by

[34] Under the low-input/traditional management scenario, farming is predominantly subsistence-based, with practices involving the use of traditional cultivars, labor-intensive techniques, minimal nutrient and pest control, and limited conservation measures. Conversely, the high-input/advanced management scenario assumes improved high-yielding varieties, full mechanization with low labor intensity, and optimal applications of nutrients and chemical pest, disease, and weed control.

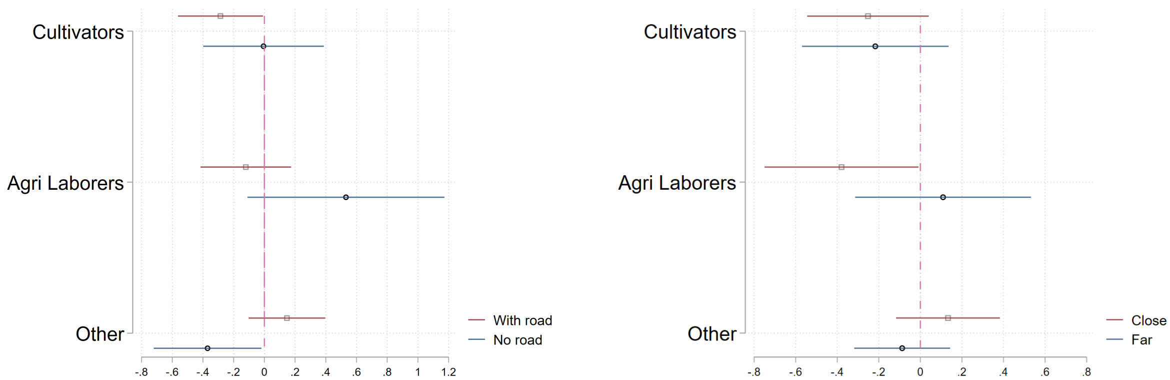
blue lines), our findings remain robust under the assumptions of high-input usage (indicated by red lines). Here, as reported in Table A.6, we document the effects on the logarithm of the number of farmers, controlling for the village population in 2001. Table A.12 confirms that our estimates are quantitatively similar when using the number of farmers as an alternative outcome variable. As a placebo test, we use data from the 2001 population census in Table A.9, revealing that the reform had no discernible effects on cultivators prior to the policy implementation for any given quartile. This indicates that the effects we found in the post-reform period are not merely a result of pre-existing trends or other factors unrelated to the bank expansion.

Labor demand or labor supply? The results presented in Figure 6 lend support to our main hypothesis that industrial growth creates more job opportunities in non-agricultural sectors and thus facilitates structural transformation. However, it is important to consider an alternative hypothesis that bank expansion may also increase industrial labor supply. Assuming non-homothetic preferences, such as a subsistence consumption requirement for agricultural goods, workers in places with lower agricultural productivity may self-select into agriculture because they are poorer (Lagakos and Waugh, 2013). Bank expansion would have a stronger impact in less productive places by relaxing this subsistence requirement and increasing industrial labor supply.

To provide additional evidence supporting our hypothesis regarding the role of labor demand, we examine the heterogeneous effects of the reform. If the labor demand hypothesis is valid, we would expect to see a stronger effect of the reform in villages that can better leverage the industrial expansion. These could include villages near urban centers (Madhok et al., 2022), where there is higher demand for labor due to industrial growth. Additionally, villages connected by roads may also experience a stronger impact, as improved infrastructure facilitates the movement of goods and labor between rural and urban areas (Asher and Novosad, 2020).

On the other hand, if the labor supply channel is the main driver of the results, we would anticipate a more pronounced response to the reform in more isolated and remote villages. In these areas, the subsistence requirement may have a greater impact on labor allocation decisions, and the relaxation of this requirement through bank expansion could result in a greater shift in labor from agriculture to non-agricultural sectors.

Figure 7 presents the impact of bank expansion on employment in villages with and without road connections (left panel), as well as those situated close to or far from population centers with populations exceeding 10,000 (right panel). It becomes evident that villages



(a) Rural Road

(b) Distance to Population Center

Figure 7: Heterogeneity in the Effect of Bank Expansion

connected by roads and those closer to population centers exhibit a stronger response to the bank expansion reform. In these locales, we observe a more pronounced decrease in the number of cultivators and agricultural laborers, coupled with an increase in 'other workers' (including workers in both manufacturing and service sectors). These patterns suggest that subsistence requirements and the labor supply channel alone are insufficient to explain the structural transformation documented in our results.

The regression results are reported in Table A.7 and A.8. Table A.13 and Table A.14 confirm that our estimates are similar using levels instead of logs as outcome variables. Using data from the 2001 population census as a placebo test, Table A.10 and Table A.11 show that the estimates tend to be much smaller and less statistically significant before the policy implementation, thus further support our identification strategy.

Intriguingly, we also note a decline in 'other workers' and a non-significant increase in agricultural laborers in villages without road connections or those situated further from population centers. This pattern suggests a spatial reallocation of agricultural activities to places with lesser access to industrial job opportunities, echoing the results reported in Madhok et al. (2022). This reallocation could be driven by two channels. First, bank expansion could directly impact agricultural investments and activities, and this effect may be more evident in remote and isolated areas. Second, there could be a general equilibrium effect of bank expansion. If farmers in neighboring villages transition to non-agricultural sectors, local food prices could increase, rendering farming more profitable. We plan to assess the relative strength of these two channels in our quantitative exercise.

6 Quantitative Model

Thus far we have presented reduced form results about the effects of bank expansion. We now discuss a quantitative model of the banking process that we hope to rationalize our results. We plan to use the model to study redistributive effects of the bank policy and consider counterfactual outcomes from alternative policies, such as direct subsidies for entering underserved areas. The quantitative model features heterogeneous firms that demand loans from banks. Banks endogeneously enter a market based on anticipated profits from selling loans.

6.1 Set Up

We consider an economy in discrete time with infinitely many periods. There is a mass Ω_t of firms in a district at each time t. Each firm lasts one period before exiting and being replaced by a new set of firms in the next period. There are N banks that offer loans to firms at time t.

6.2 Product Demand

A representative consumer purchases a consumption index given by CES aggregator

$$Y = \left(\int_{\Omega_t} Y_i^{\frac{\eta-1}{\eta}} di \right)^{\frac{\eta}{\eta-1}}$$

Y_i is the amount of output produced by firm i.

6.3 Firm Problem

The firm's problem proceeds in two stages. In the first stage the firm decides whether to seek financing. In the second stage the firm makes its production decision. Since firms live for only a single period we omit time subscripts unless necessary.

6.3.1 Stage 1: Borrowing

First, each firm i decides whether to take out a loan to start production. Firm i has productivity $A_i \in \{0, \bar{A}\}$, which is unknown in Stage 1 and follows a Bernoulli distribution with probability of success σ_i. We assume all agents observe σ_i, which represents public information about the firm and is drawn iid from distribution function $F_{\sigma,t}$, which can vary with the time period, t. A firm i that wishes to produce pays an entry cost c. It then chooses a single bank j to borrow from and takes out a loan of size M_{ijt}. Alternatively the firm can choose an outside option, in which case the firm does not take out a loan and does not proceed to the second stage.

6.3.2 Stage 2: Production

At the start of Stage 2, i realizes its draw of A_i. If $A_i = 0$, i declares bankruptcy and exits with 0 payoff. If $A_i = \bar{A}$, then firm i has a Cobb-Douglas production function given by

$$Y(L, K) = \bar{A} L^{\alpha} K^{1-\alpha} \tag{1}$$

The firm uses its loan M_{ijt} to hire labor and rent capital, facing the budget constraint

$$w_t L + r_t K \leq M_{ijt} \tag{2}$$

where w_t is the wage rate and r_t is the rental rate of capital at time t.

6.4 Firm Payoffs

We now describe the firm's choice problem and payoffs given the structure in Stage 1 and Stage 2. In the second stage, payoffs from production are

$$\Pi^p_{it}(M_{ijt}) \equiv \max_{L,K} \sigma_i \big(p(Y(L,K))Y(L,K) - w_t L - r_t K\big) \tag{3}$$

$$\text{s.t. } w_t L + r_t K \leq M_{ijt}$$

The firm produces with probability σ_i, in which case it earns the difference between its revenue and cost of inputs. The price function $p(Y)$ follows from CES demand for products. Input choices depend on the loan size, M_{ijt}. If the firm does not produce it exits and earns 0 profits in production.

In the first stage, i chooses whether to borrow, which bank to borrow from, and how much to borrow. Payoffs are determined by

$$\Pi_{it} \equiv \max\{\epsilon_{io}, \max_j \max_M \big(\Pi^p_{it}(M) - R_{ijt}M + \epsilon_{ij}\big) - c\} \tag{4}$$

The maximization over M determines i's choice of how much to borrow, which depends on the interest rates offered by banks, R_{ijt}. The maximization over j determines i's choice of which bank to borrow from. We assume that ϵ_{ij} is a Type 1 Extreme Value distributed random variable, reflecting firm preferences to borrow from a given bank j. These preferences could represent distance costs or other idiosyncratic factors that push a firm to choose one bank over another for reasons other than the interest rate. The outermost maximization determines whether i chooses the outside option, which gives a payoff determined by the Type 1 Extreme Value shock ϵ_{io}. c is an entry cost that firms need to pay in order to produce.

6.5 Bank Static Problem

We now describe each bank's static period-by-period maximization problem before discussing the dynamic entry problem. We assume that each bank in the market at time t is identical aside from the vector of firm preference shocks, $\vec{\epsilon}_j$, which are private information of the firms. The bank's problem is to choose an interest rate to offer in each period to each firm. Banks may condition the interest rate on the observed characteristics of the firm, σ_i.

Given logit preference shocks for firms, a bank j that sets interest rate R_{ijt} for firm i will sell the loan with probability

$$p_{ijt}(R_{ijt}) = \frac{\exp\Big(-\gamma\left(\Pi^p_{it}(M_{ijt}(R_{ijt})) - R_{ijt}M_{ijt}(R_{ijt}) - c\right)\Big)}{1 + \sum_{j' \in \mathcal{B}_t} \exp\Big(-\gamma\left(\Pi^p_{it}(M_{ij'}(R_{ij't})) - R_{ij't}M_{ij't}(R_{ij't}) - c\right)\Big)} \tag{5}$$

where γ is the scale parameter for the logit shocks and loan sizes M_{ijt} depend endogeneously on the interest rate. $\mathcal{B}_t$ is the set of banks operating at time t. The expected payoff for the bank is thus

$$\Pi^b_{ijt} \equiv \max_{R_{ijt}} p_{ijt}(R_{ijt})\left[\sigma_i R_{ijt}M_{ijt}(R_{ijt}) - (1-\sigma_i)M_{ijt}(R_{ijt})\right] \tag{6}$$

If the bank sets an interest rate R_{ijt} it sells the loan with probability $p_{ijt}(R_{ijt})$. If it does sell the loan, the loan gets repaid with interest with probability σ_i. If the firm realizes a low productivity draw the bank loses the entirety of the loan amount. In each period the bank chooses an interest rate to maximize Equation 6 for each firm.

6.6 Bank Dynamic Problem

In each period, the set of banks is composed of incumbent banks and entrant banks. Dynamic payoffs for incumbent banks are the present discounted value of future static payoffs:

$$\Pi_t^* \equiv \sum_{s=t}^{\infty} \beta^s \Omega_s \int \Pi_{ijs}^b dF_{\sigma,s} \tag{7}$$

where β is the discount rate. Entrant banks have to pay a fixed cost δ in order to enter. Their payoffs are

$$\Pi_t^E \equiv \Pi_t^* - \delta. \tag{8}$$

6.7 Dynamic Equilibrium

An equilibrium of this economy is one in which at each time period firms choose to take loans from their preferred bank, or not enter; firms produce optimally given their loan size; banks set interest rates optimally; and the number of banks at each time period obeys a zero profit condition for potential entrants.

6.8 Characteristics of Equilibrium

Because banks are symmetric, they each offer the same interest rate to a firm i at time t. The interest rate offered at time t depends on the firm's type probability. In turn the firm's loan size depends on σ_i, the offered interest rate, and the prevailing wage and rental rate of capital.

Equation 7 shows that bank profits depend on market size, Ω_t, and the distribution of σ, $F_{\sigma,t}$. For entrants, profits also depend on the entry cost, δ. From Equations 5 and 6, profits also depend on the number of other banks in the market. Since banks are symmetric, the number of banks determines the degree of competition in the banking sector. When the number of banks is small, banks' pricing decisions have equilibrium effects on total market size, resulting in higher interest rates. As the number of banks increases interest rates fall. In turn firms that face lower interest rates take out bigger loans, which increases their labor inputs, capital inputs, output, and revenue.

6.9 Model Predictions

The model predicts that the key features for understanding bank entry decisions are the market size, Ω, the number of incumbent banks, N, the distribution of firm types, F_σ, and the entry cost, δ. Areas with growing populations and improved firm creditworthiness over time should experience growth in the number of bank branches. Remaining cross-district variation in bank presence comes from differences in entry costs. As the number of banks increases firm sizes should also increase due to lower interest rates. These predictions match our empirical findings in Section 4.

Our model also accommodates endogeneous entry into production and endogenous choices for whether to borrow from a formal lending institution. Firms with very low probability of being a high type choose not produce because of the entry cost. Firms with intermediate probability of being a high type may choose to borrow from an informal lending institution whose loans are not reported in the data. We can incorporate an informal lender in our framework by having an additional lender in the set of banks, $\mathcal{B}_t$, with the informal lender's interest rate exogeneously fixed.

Further, bank entry can have important distributional effects on firms. Lower interest rates offered by banks may benefit firms with high σ_i more than firms with low σ_i. The reason is that firms with low σ_i may not enter at all anyway, or if they do enter, may be more likely to choose informal lending. Because of CES demand, expansion by high σ_i firms may crowd out product demand for low σ_i firms. Such patterns may explain our empirical results regarding heterogeneous treatment effects across firms.

6.10 Calibration

The important model objects to take to the data are Ω, N, F_σ, and δ. District market size and the number of incumbent banks can be measured using census records and administrative data from the Reserve Bank of India. To estimate the distribution of firm types, we use Equations 5 and 6, which implicitly relate variation in firm types to variation in interest rates when there are many banks. We infer the distribution of firm interest rates in a district using firm survey data from Prowess. In taking the model to the data, we also allow for a district-level time-invariant shifter that subsumes the entry cost and other parameters fixed over time. The shifter reflects variation in the institutional environment of a district that makes it more or less attractive for banks to enter. We calibrate this parameter to explain residual variation in bank entry rates after accounting for Ω, N and F_σ.

6.11 Policy Counterfactuals

The model allows us to simulate the effects of the bank expansion policy and other counterfactuals. The policy reform we study has a natural analogue to the model objects. "Under-banked" districts are ones in which the population to bank branch ratio, Ω/N, is above the national average. The policy changes banks' incentives over which districts to enter. In particular, a potential entrant can enter a market with low Ω/N only if it also enters a market with relatively high Ω/N. The policy thus has redistributive effects across districts, allowing banks to enter districts with strong institutional environments only if they also enter markets with less favorable environments. Using the model, we can simulate which districts would have experienced entry absent the policy given changes in Ω/N and F_σ. We can then compare the counterfactual outcomes with the observed outcomes to determine the redistributive effects of the policy. We can also consider alternative policies such as direct entry subsidies targeted to districts with poor financial access. These counterfactuals will be informative about bank policy in India and other developing countries with weaker financial institutions and regional inequality.

6.12 Extensions to the Model

An important extension to the model is to specify a labor market and preferences for workers. Adding these features can allow us to speak to the empirical results on structural transformation from Section 5. It will also enable us to analyze the effects of the policy on welfare.

7 Concluding Remarks

In conclusion, this paper provides a comprehensive analysis of the impacts of bank expansion on industrial growth and labor allocation based on a nationwide policy experiment in India. Our findings reveal that the expansion of banks in under-banked districts alleviated firms' borrowing constraints, stimulating their growth in sales revenue, employment, and capital accumulation. However, contrary to initial expectations, the expansion did not translate into increased firm dynamics or higher entry rates. Moreover, we find that this policy had minimal effects on product innovation and total factor productivity. This policy primarily benefited incumbent firms and left potential entrants largely unaffected, hinting at limited dynamic gains from bank expansion.

On the labor front, we find that bank expansion created more high-paying jobs and led to a significant labor reallocation towards the manufacturing sector, particularly in regions with lower agricultural productivity. This labor movement underscores the role of manufacturing development as a significant "labor pull" factor, shaping the process of structural transformation.

References

Allcott, Hunt, Allan Collard-Wexler, and Stephen D O'Connell, "How do electricity shortages affect industry? Evidence from India," *American Economic Review*, 2016, *106* (3), 587–624.

Alvarez-Cuadrado, Francisco and Markus Poschke, "Structural change out of agriculture: Labor push versus labor pull," *American Economic Journal: Macroeconomics*, 2011, *3* (3), 127–158.

Asher, Sam and Paul Novosad, "Rural Roads and Local Economic Development," *American Economic Review*, March 2020, *110* (3), 797–823.

Banerjee, Abhijit, Dean Karlan, and Jonathan Zinman, "Six Randomized Evaluations of Microcredit: Introduction and Further Steps," *American Economic Journal: Applied Economics*, January 2015, *7* (1), 1–21.

Banerjee, Abhijit V. and Esther Duflo, "Do Firms Want to Borrow More? Testing Credit Constraints Using a Directed Lending Program," *The Review of Economic Studies*, April 2014, *81* (2), 572–607.

Baqaee, David Rezza and Emmanuel Farhi, "Productivity and Misallocation in General Equilibrium*," *The Quarterly Journal of Economics*, February 2020, *135* (1), 105–163.

Barboni, Giorgia, Erica Field, and Rohini Pande, "Rural Banks Can Reduce Poverty: Evidence from 870 Indian Villages," 2021.

Bartelsman, Eric, John Haltiwanger, and Stefano Scarpetta, "Cross-Country Differences in Productivity: The Role of Allocation and Selection," *American Economic Review*, February 2013, *103* (1), 305–334.

Bau, Natalie and Adrien Matray, "Misallocation and Capital Market Integration: Evidence From India," *Econometrica*, 2023, *91* (1), 67–106. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.3982/ECTA19039.

Bazzi, Samuel, Raquel de Freitas Oliveira, Marc-A Muendler, and James Rauch, "Credit Supply Shocks and Firm Dynamics: Evidence from Brazil," 2021, p. 50.

Bencivenga, Valerie R. and Bruce D. Smith, "Unemployment, migration, and growth," *Journal of Political Economy*, 1997, *105* (3), 582–608.

Bertrand, Marianne, Chang-Tai Hsieh, and Nick Tsivanidis, "Contract Labor and Firm Growth in India," August 2021.

Breza, Emily and Cynthia Kinnan, "Measuring the Equilibrium Impacts of Credit: Evidence from the Indian Microfinance Crisis*," *The Quarterly Journal of Economics*, August 2021, *136* (3), 1447–1497.

Bruhn, Miriam and Inessa Love, "The Real Impact of Improved Access to Finance: Evidence from Mexico," *The Journal of Finance*, 2014, *69* (3), 1347–1376. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/jofi.12091.

Buera, Francisco J., Joseph P. Kaboski, and Yongseok Shin, "Finance and Development: A Tale of Two Sectors," *American Economic Review*, August 2011, *101* (5), 1964–2002.

Burgess, Robin and Rohini Pande, "Do Rural Banks Matter? Evidence from the Indian Social Banking Experiment," *American Economic Review*, June 2005, *95* (3), 780–795.

Bustos, Paula, Bruno Caprettini, and Jacopo Ponticelli, "Agricultural Productivity and Structural Transformation: Evidence from Brazil," *American Economic Review*, June 2016, *106* (6), 1320–1365.

Calonico, Sebastian, Matias D. Cattaneo, and Rocio Titiunik, "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs," *Econometrica*, 2014, *82* (6), 2295–2326. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.3982/ECTA11757.

Cramer, Kim Fe, "Bank Presence and Health," SSRN Scholarly Paper 3917526, Social Science Research Network, Rochester, NY September 2021.

David, Joel M. and Venky Venkateswaran, "The Sources of Capital Misallocation," *American Economic Review*, July 2019, *109* (7), 2531–2567.

Demirguc-Kunt, Asli, Leora Klapper, Dorothe Singer, Saniya Ansar, and Jake Hess, *Global Findex Database 2017: Measuring Financial Inclusion and the Fintech Revolution*, Washington, DC: World Bank, April 2018.

Donaldson, Dave and Adam Storeygard, "The View from Above: Applications of Satellite Data in Economics," *Journal of Economic Perspectives*, November 2016, *30* (4), 171–198.

Egger, Dennis, Johannes Haushofer, Edward Miguel, Paul Niehaus, and Michael Walker, "General Equilibrium Effects of Cash Transfers: Experimental Evidence From Kenya," *Econometrica*, 2022, *90* (6), 2603–2643. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.3982/ECTA17945.

FDIC, "2021 FDIC National Survey of Unbanked and Underbanked Households," 2021.

Fonseca, Julia and Adrien Matray, "The Real Effects of Banking the Poor: Evidence from Brazil," 2022.

Garcia-Macia, Daniel, Chang-Tai Hsieh, and Peter J. Klenow, "How Destructive Is Innovation?," *Econometrica*, 2019, *87* (5), 1507–1541. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.3982/ECTA14930.

Gollin, Douglas, David Lagakos, and Michael E. Waugh, "The Agricultural Productivity Gap *," *The Quarterly Journal of Economics*, May 2014, *129* (2), 939–993.

_ , Stephen Parente, and Richard Rogerson, "The Role of Agriculture in Development," *American Economic Review*, 2002, *92* (2), 160–164.

Grembi, Veronica, Tommaso Nannicini, and Ugo Troiano, "Do Fiscal Rules Matter?," *American Economic Journal: Applied Economics*, July 2016, *8* (3), 1–30.

Gylfason, Thorvaldur and Gylfi Zoega, *The road from agriculture*, Cambridge, MA: MIT Press, 2006.

Harris, John R. and Michael P. Todaro, "Migration, Unemployment and Development: 2-Sector Analysis," *American Economic Review*, 1970, *60* (1), 126–142.

Harrison, Ann, Benjamin Hyman, Leslie Martin, and Shanthi Nataraj, "When do Firms Go Green? Comparing Command and Control Regulations with Price Incentives in India," November 2015.

Hsieh, Chang-Tai and Peter J. Klenow, "Misallocation and Manufacturing TFP in China and India," *Q J Econ*, November 2009, *124* (4), 1403–1448. Publisher: Oxford Academic.

_ and _ , "The Life Cycle of Plants in India and Mexico *," *The Quarterly Journal of Economics*, August 2014, *129* (3), 1035–1084.

Ji, Yan, Songyuan Teng, and Robert M. Townsend, "Dynamic Bank Expansion: Spatial Growth, Financial Access, and Inequality," *Journal of Political Economy*, February 2023. Publisher: The University of Chicago Press.

Kerr, William R. and Ramana Nanda, "Democratizing entry: Banking deregulations, financing constraints, and entrepreneurship," *Journal of Financial Economics*, October 2009, *94* (1), 124–149.

Kone, Zovanga L, Maggie Y Liu, Aaditya Mattoo, Caglar Ozden, and Siddharth Sharma, "Internal borders and migration in India*," *Journal of Economic Geography*, July 2018, *18* (4), 729–759.

Lagakos, David and Michael E. Waugh, "Selection, Agriculture, and Cross-Country Productivity Differences," *American Economic Review*, April 2013, *103* (2), 948–980.

Lee, David S. and Thomas Lemieux, "Regression Discontinuity Designs in Economics," *Journal of Economic Literature*, June 2010, *48* (2), 281–355.

Levinsohn, James and Amil Petrin, "Estimating Production Functions Using Inputs to Control for Unobservables," *The Review of Economic Studies*, 04 2003, *70* (2), 317–341.

Lewis, William Arthur, "Economic Development with Unlimited Supplies of Labour," *The Manchester School of Economic and Social Studies*, 1954, *22* (2), 139–191.

Li, Nicholas, "An Engel Curve for Variety," *The Review of Economics and Statistics*, March 2021, *103* (1), 72–87.

Madhok, Raahil, Frederik Noack, Ahmed Mushfiq Mobarak, and Olivier Deschenes, "Rural-Urban Migration and The Re-organization of Agriculture," *Discussion Papers*, October 2022.

Martin, Leslie A., Shanthi Nataraj, and Ann E. Harrison, "In with the Big, Out with the Small: Removing Small-Scale Reservations in India," *The American Economic Review*, 2017, *107* (2), 354–386. Publisher: American Economic Association.

McCrary, Justin, "Manipulation of the running variable in the regression discontinuity design: A density test," *Journal of Econometrics*, February 2008, *142* (2), 698–714.

Melitz, Marc J., "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity," *Econometrica*, 2003, *71* (6), 1695–1725. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/1468-0262.00467.

Munshi, Kaivan and Mark Rosenzweig, "Networks and Misallocation: Insurance, Migration, and the Rural-Urban Wage Gap," *American Economic Review*, January 2016, *106* (1), 46–98.

Ngai, L. Rachel and Christopher A. Pissarides, "Structural Change in a Multisector Model of Growth," *American Economic Review*, 2007, *97* (1), 429–443.

Porzio, Tommaso, Federico Rossi, and Gabriella Santangelo, "The Human Side of Structural Transformation," *American Economic Review*, 2022, *112* (8), 2774–2814.

Restuccia, Diego and Richard Rogerson, "The Causes and Costs of Misallocation," *Journal of Economic Perspectives*, August 2017, *31* (3), 151–174.

Sraer, David and David Thesmar, "A Sufficient Statistics Approach for Aggregating Firm-Level Experiments," January 2018.

Young, Alwyn, "Structural Transformation, the Mismeasurement of Productivity Growth, and the Cost Disease of Services," *American Economic Review*, November 2014, *104* (11), 3635–3667.

Young, Nathaniel, "Banking and Growth: Evidence From a Regression Discontinuity Analysis," SSRN Scholarly Paper 3104632, Social Science Research Network, Rochester, NY November 2017.

A Empirical Appendix

A.1 Empirical Appendix

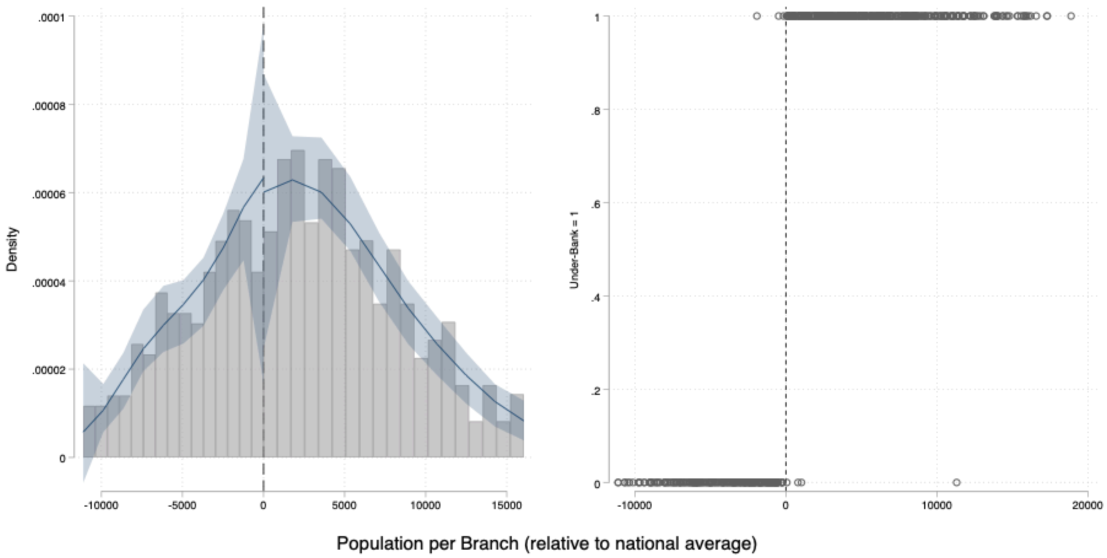
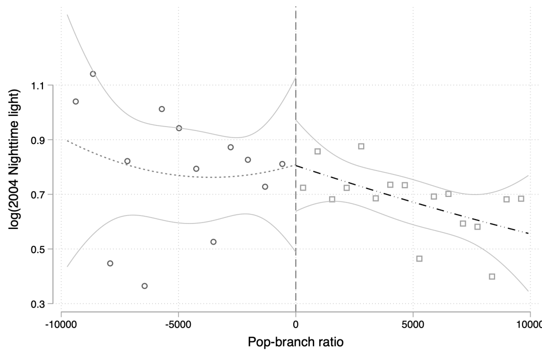
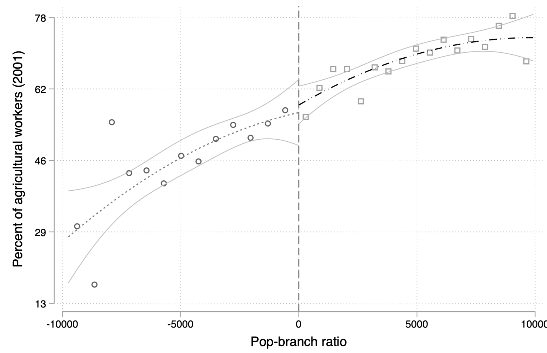


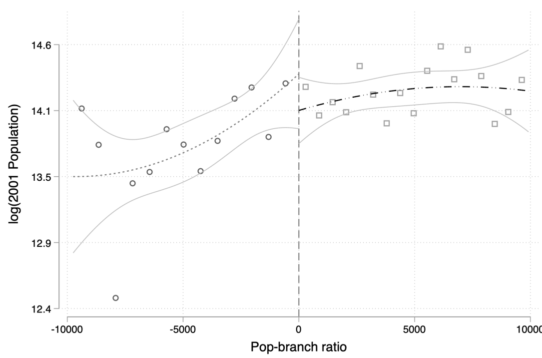
Figure A.1: MaCrary manipulation test and the first stage.



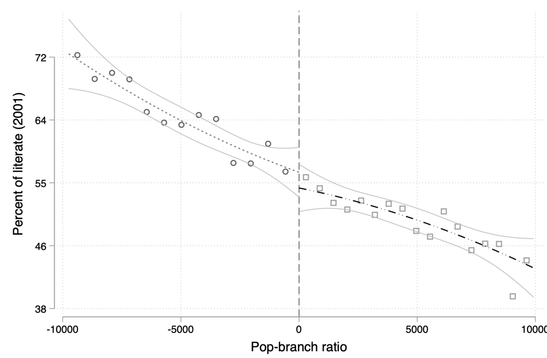
(a) log(Light intensity) in 2004



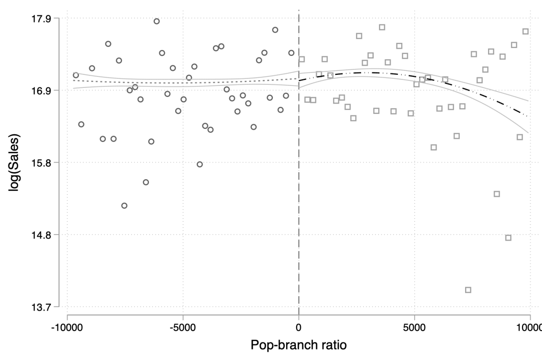
(b) Percent of agricultural workers



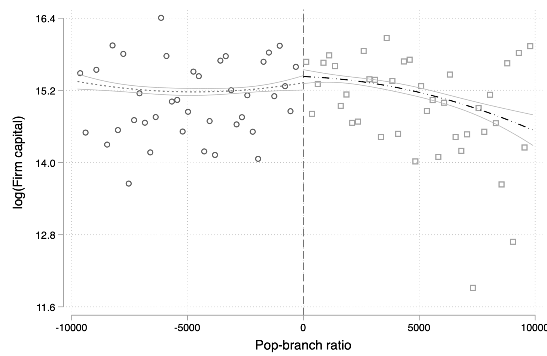
(c) log(Population)



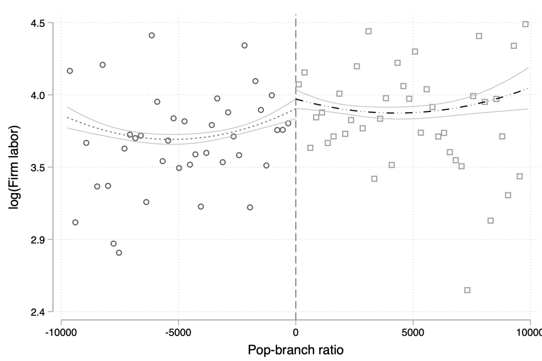
(d) Percent literate



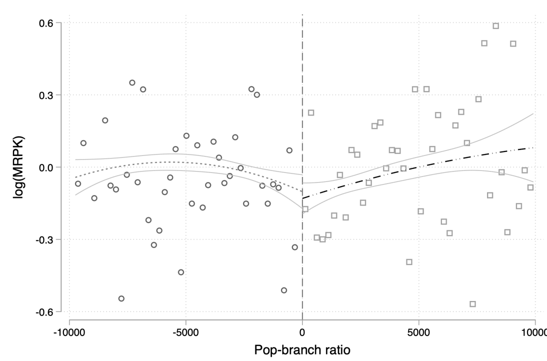
(e) log(Firm sales)



(f) log(Firm fixed asset)

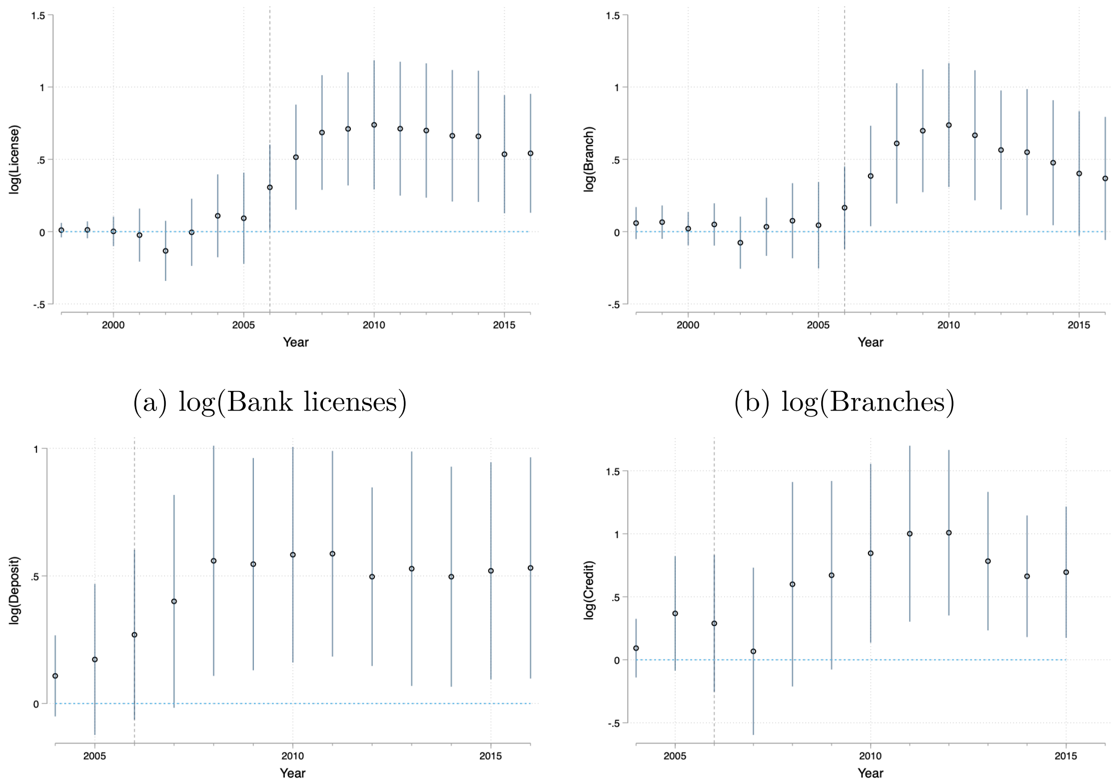


(g) log(Firm employee)



(h) log(Firm MRPK)

Figure A.2: Smoothness of pre-policy covariates



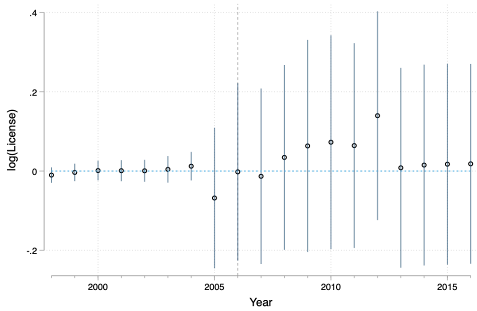
(a) log(Bank licenses)

(b) log(Branches)

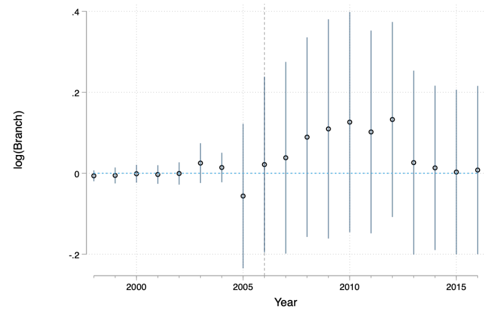
(c) log(Deposit)

(d) log(Credit)

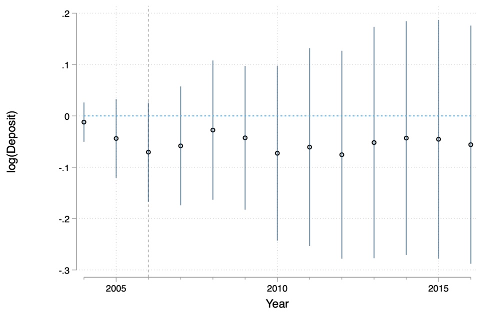
Figure A.3: Efffects on Private Sector Banks



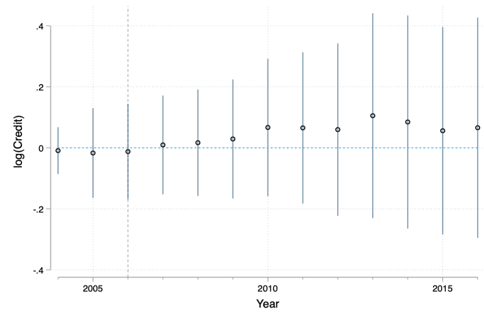
(a) log(Public Bank licenses)



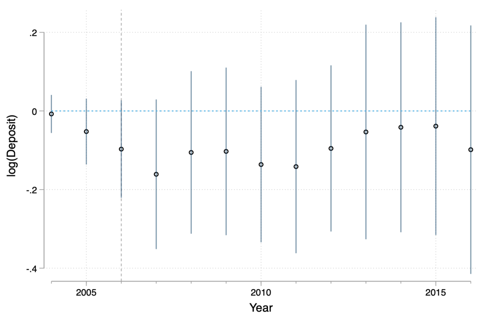
(b) log(Public Bank Branches)



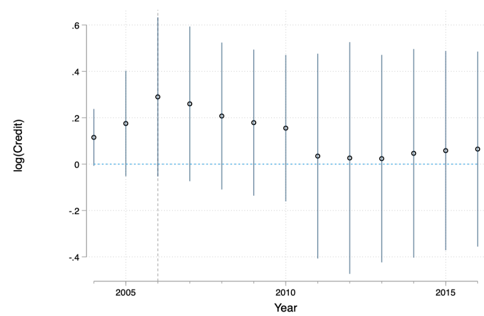
(c) log(Nationalized Bank Deposit)



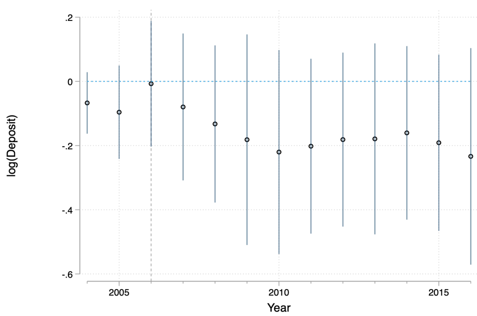
(d) log(Nationalized Bank Credit)



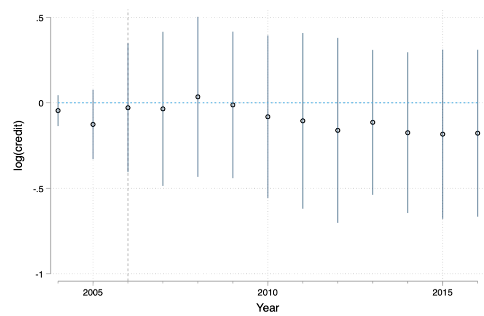
(e) log(SBI Deposit)



(f) log(SBI Credit)



(g) log(RRB Deposit)



(h) log(RRB Credit)

Figure A.4: Effects on Public Sector Banks

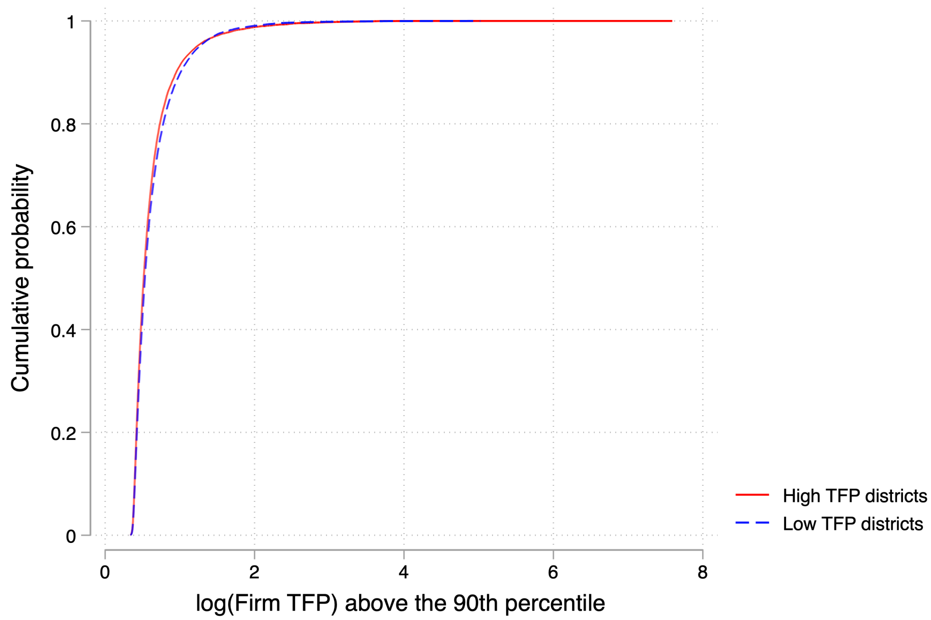


Figure A.5: Truncated Distribution of log(TFP) for district with Low and High Aggregate TFP

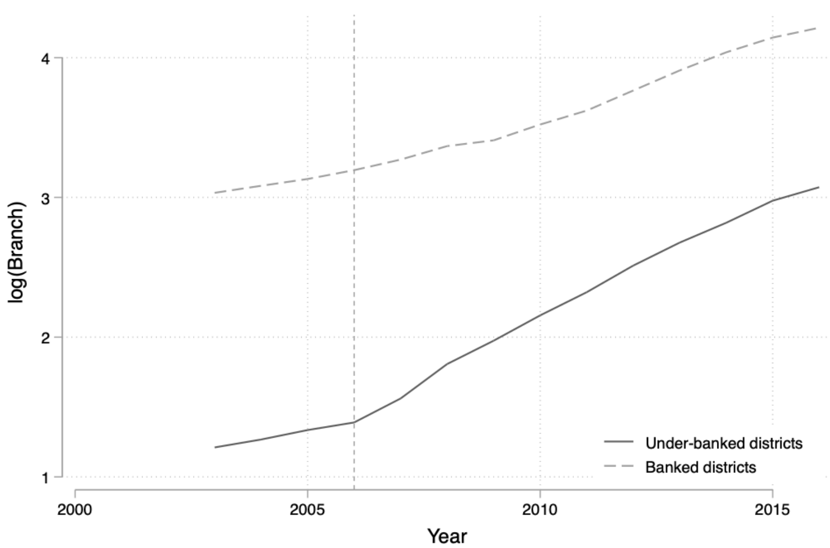
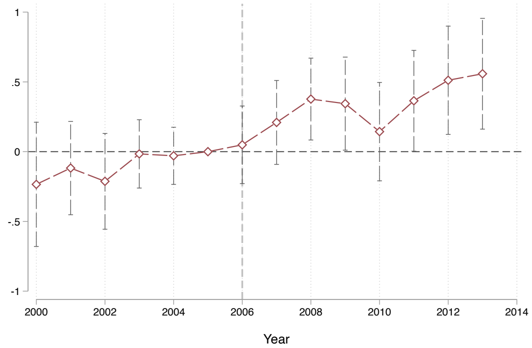
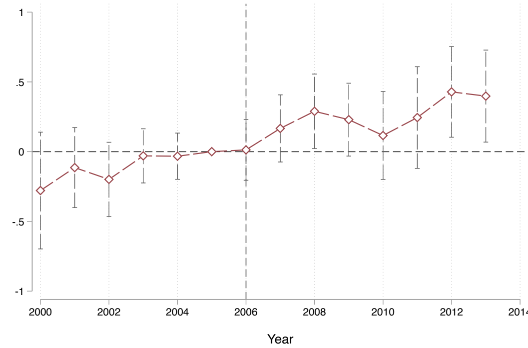


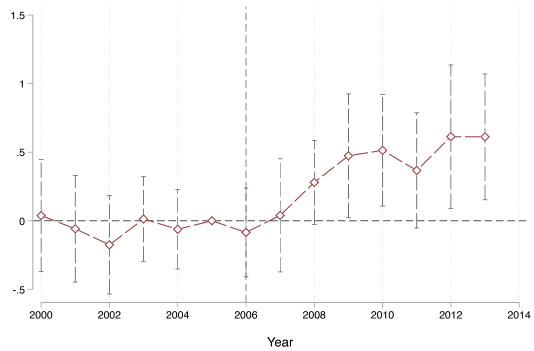
Figure A.6: Number of branches in treated versus control districts





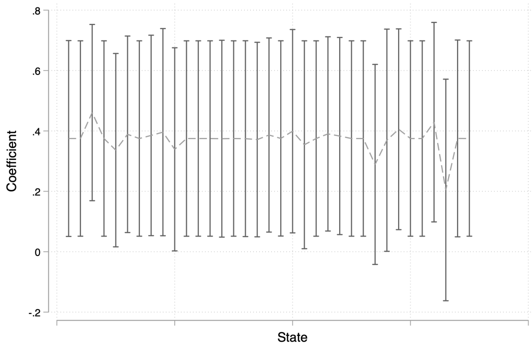
(a) log(Wage Bill)

(b) log(Employment)

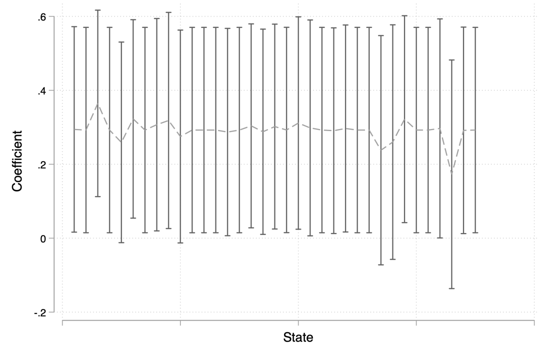


(c) log(Outstanding Loan)

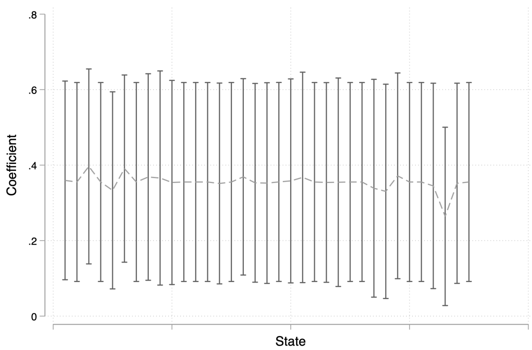
Figure A.7: Event Study Graphs for the Treatment Effects on Wage Bill, Employment and Outstanding Loan



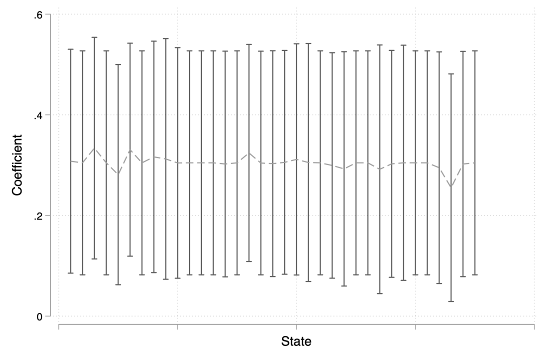
(a) log(Fixed Assets)



(b) log(Sales Revenue)

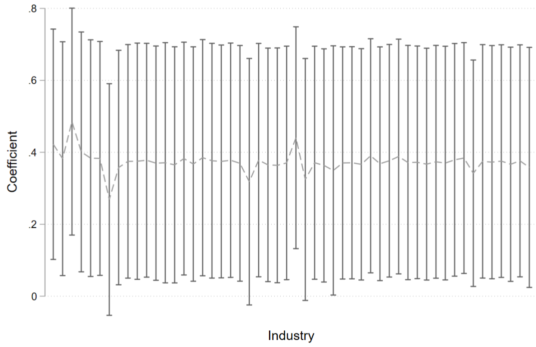


(c) log(Wage Bills)

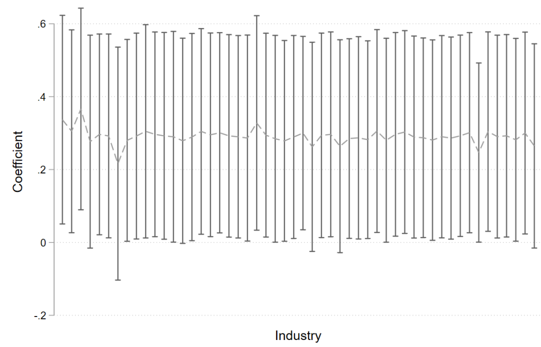


(d) log(Employment)

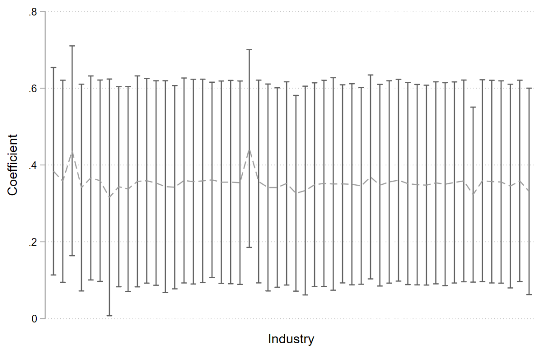
Figure A.8: Robustness: Drop Individual States



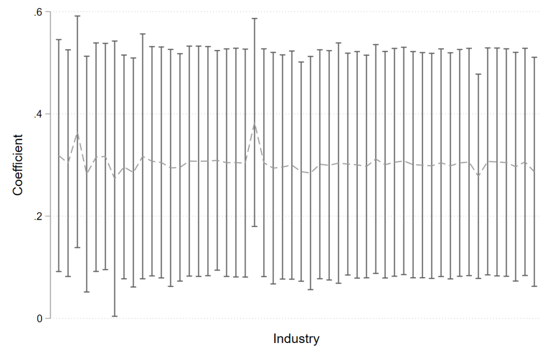
(a) log(Fixed Assets)



(b) log(Sales Revenue)

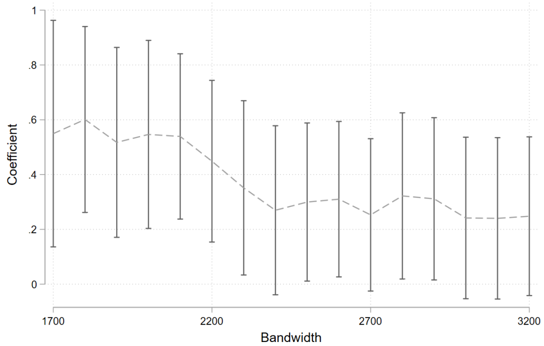


(c) log(Wage Bills)

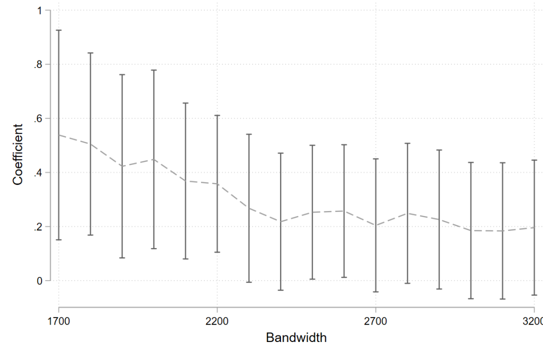


(d) log(Employment)

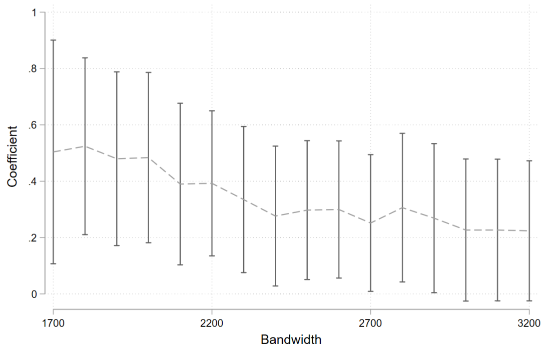
Figure A.9: Robustness: Drop Individual 2-digit Industries



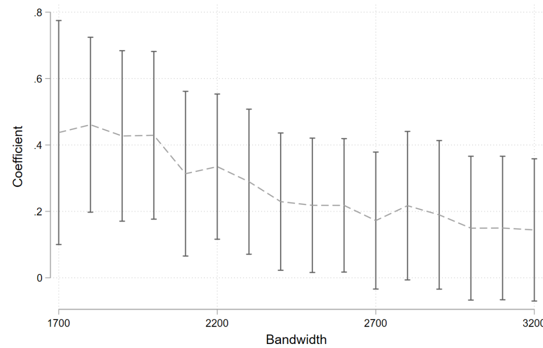
(a) log(Fixed Assets)



(b) log(Sales Revenue)

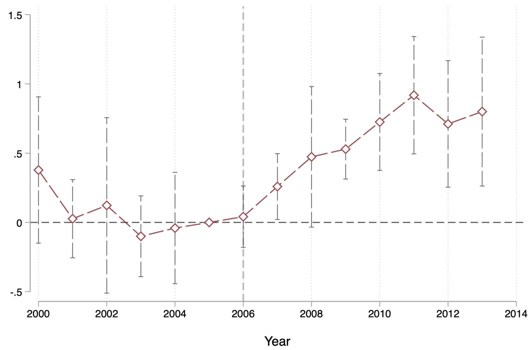


(c) log(Wage Bills)

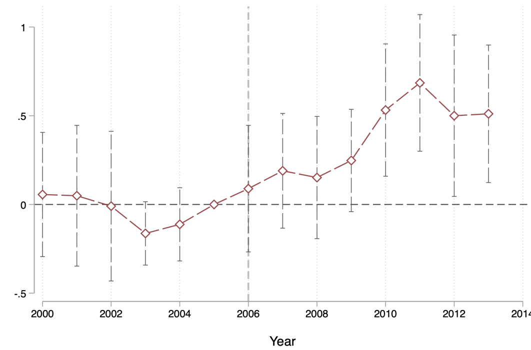


(d) log(Employment)

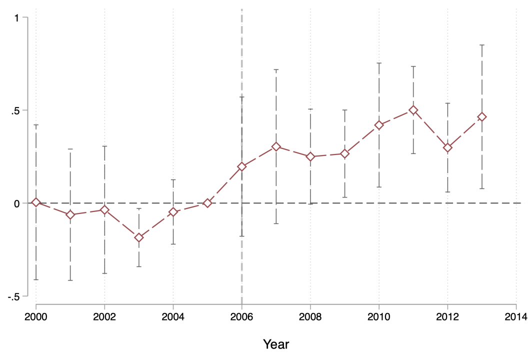
Figure A.10: Robustness: Change Bandwidth



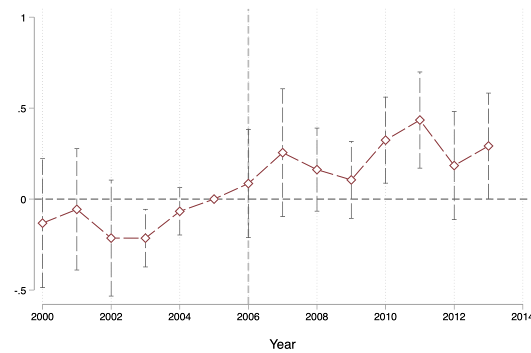
(a) log(Fixed Assets)



(b) log(Sales Revenue)



(c) log(Wage Bills)



(d) log(Employment)

Figure A.11: District Aggregate Outcomes: Event Study Plots

Table A.1: Treatment Effect of Bank Expansion on Firms: Parsimonious Specification

	(1)	(2)	(3)	(4)
Dependent Variable	Revenues	Capital	Wages	Employment
Treated * Post	0.322**	0.407**	0.395**	0.334**
	(0.159)	(0.180)	(0.157)	(0.130)
Observations	135,797	135,797	135,797	135,797
R-squared	0.146	0.174	0.161	0.110
District FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes

Notes: All outcome variables are in logs. Standard errors are clustered at the district level. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.2: Treatment Effect of Bank Expansion on Firms: High-Dimensional Fixed Effects

	(1)	(2)	(3)	(4)
Dependent Variable	Revenues	Capital	Wage Bills	Employment
Treated * Post	0.270***	0.300***	0.317***	0.253***
	(0.094)	(0.099)	(0.096)	(0.081)
Observations	134,929	134,929	134,929	134,929
R-squared	0.262	0.283	0.262	0.227
District FE	Yes	Yes	Yes	Yes
Industry*Year FE	Yes	Yes	Yes	Yes
State*Year FE	Yes	Yes	Yes	Yes
Firm Controls	Yes	Yes	Yes	Yes
District Trends	Yes	Yes	Yes	Yes

Notes: All outcome variables are in logs. State*Year FE are Indian states interacted with year fixed effects. Industry*Year FE are 2-digit industry interacted with year fixed effects. Firm Controls include firm ownership fixed effects and a dummy variable of being in urban areas. District Trends include district population in 2001 and the number of bank branches in 1997, interacted with a linear time trend. Standard errors are clustered at the district level. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.3: Treatment Effect of Bank Expansion on Capital

	(1)	(2)	(3)	(4)
Dependent Variable	Land	Building	Plant/Machine	Computer
Treated * Post	0.285**	0.481***	0.648**	0.412
	(0.143)	(0.173)	(0.309)	(0.254)
Observations	100,387	121,476	133,200	88,237
R-squared	0.173	0.192	0.198	0.120
District FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Firm Controls	Yes	Yes	Yes	Yes
District Trends	Yes	Yes	Yes	Yes

Notes: All outcome variables are in logs. All capital variables are deflated to constant 2004-2005 Rupee using the Gross Capital Formation data from the RBI. Firm Controls include firm ownership fixed effects and a dummy variable of being in urban areas. District Trends include district population in 2001 and the number of bank branches in 1997, interacted with a linear time trend. Standard errors are clustered at the district level. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.4: Treatment Effect of Bank Expansion on Outstanding Loan

Dependent Variable	Outstanding Loan			$\mathbf{1}$(Loan > 0)		
	(1)	(2)	(3)	(4)	(5)	(6)
Treated*Post	0.415**	0.465**	0.244**	0.016	0.012	0.003
	(0.171)	(0.184)	(0.119)	(0.022)	(0.021)	(0.020)
Observations	109,832	109,914	109,241	135,673	135,797	134,929
R-squared	0.129	0.093	0.195	0.068	0.064	0.105
District FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry*Year FE	No	No	Yes	No	No	Yes
State*Year FE	No	No	Yes	No	No	Yes
Firm Controls	Yes	No	Yes	Yes	No	Yes
District Trends	Yes	No	Yes	Yes	No	Yes

Notes: Outstanding Loan variables are in logs. $\mathbf{1}$(Loan > 0) is a dummy variable equal to 1 if the firm reports a positive outstanding loan in a year. State*Year FE are Indian states interacted with year fixed effects. Industry*Year FE are 2-digit industry interacted with year fixed effects. Firm Controls include firm ownership fixed effects and a dummy variable of being in urban areas. District Trends include district population in 2001 and the number of bank branches in 1997, interacted with a linear time trend. Standard errors are clustered at the district level. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.5: Treatment Effect of Bank Expansion on Firm Productivity

	TFP					
	OLS			LP		
	(1)	(2)	(3)	(4)	(5)	(6)
Treated*Post	0.023	0.024	0.009	0.035	0.037	0.022
	(0.022)	(0.022)	(0.016)	(0.026)	(0.026)	(0.017)
Observations	135,509	135,633	134,770	135,509	135,633	134,770
R-squared	0.029	0.027	0.105	0.037	0.034	0.098
District FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	No	Yes	Yes	No
Industry*Year FE	No	No	Yes	No	No	Yes
State*Year FE	No	No	Yes	No	No	Yes
Firm Controls	Yes	No	Yes	Yes	No	Yes
District Trends	Yes	No	Yes	Yes	No	Yes

Notes: All outcome variables are in logs. TFP is measured by the OLS residuals (columns 1-3) and estimating revenue production functions (columns 4-6) following Levinsohn and Petrin (2003). State*Year FE are Indian states interacted with year fixed effects. Industry*Year FE are 2-digit industry interacted with year fixed effects. Firm Controls include firm ownership fixed effects and a dummy variable of being in urban areas. District Trends include district population in 2001 and the number of bank branches in 1997, interacted with a linear time trend. Standard errors are clustered at the district level. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.6: Treatment Effect of Bank Expansion on Farmers by Crop Suitability Quartiles

	Panel A: Crop Suitability, High Input Use			
	(1) Quartile 1	(2) Quartile 2	(3) Quartile 3	(4) Quartile 4
Conventional	-0.377***	-0.331**	-0.240*	-0.151
	(0.135)	(0.154)	(0.124)	(0.135)
Bias-corrected	-0.428***	-0.387**	-0.282**	-0.185
	(0.135)	(0.154)	(0.124)	(0.135)
Robust	-0.428***	-0.387**	-0.282**	-0.185
	(0.155)	(0.172)	(0.143)	(0.151)
Observations	179,465	117,043	107,733	74,178
State FE	Yes	Yes	Yes	Yes
Village Controls	Yes	Yes	Yes	Yes

	Panel B: Crop Suitability, Low Input Use			
	(1) Quartile 1	(2) Quartile 2	(3) Quartile 3	(4) Quartile 4
Conventional	-0.392***	-0.234	-0.239	-0.095
	(0.131)	(0.153)	(0.154)	(0.141)
Bias-corrected	-0.444***	-0.282*	-0.277*	-0.135
	(0.131)	(0.153)	(0.154)	(0.141)
Robust	-0.444***	-0.282	-0.277	-0.135
	(0.152)	(0.174)	(0.172)	(0.160)
Observations	162,389	119,715	114,646	82,845
State FE	Yes	Yes	Yes	Yes
Village Controls	Yes	Yes	Yes	Yes

Notes: The dependent variables are expressed in logarithmic form. Standard errors clustered at the district level in parentheses. Crop suitability refers to cereal crop potential production measure (low/high input usage, log) from the FAO Global Agro-Ecological Zones (GAEZ) at the village level. Villages are then sorted into quartiles based on their crop suitability within a sub-district.* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.7: Treatment Effect of Bank Expansion on Employment by Road

Panel A: With Road

	(1) Cultivators	(2) Agri. Laborers	(3) Other Workers
Conventional	-0.286**	-0.121	0.147
	(0.141)	(0.151)	(0.127)
Bias-corrected	-0.331**	-0.151	0.180
	(0.141)	(0.151)	(0.127)
Robust	-0.331**	-0.151	0.180
	(0.163)	(0.175)	(0.155)
Observations	314,962	286,293	313,112
State FE	Yes	Yes	Yes
Village Controls	Yes	Yes	Yes

Panel B: Without Road

	(1) Cultivators	(2) Agri. Laborers	(3) Other Workers
Conventional	-0.006	0.531	-0.370**
	(0.201)	(0.327)	(0.179)
Bias-corrected	-0.032	0.556*	-0.385**
	(0.201)	(0.327)	(0.179)
Robust	-0.032	0.556	-0.385*
	(0.257)	(0.438)	(0.225)
Observations	165,804	127,494	156,479
State FE	Yes	Yes	Yes
Village Controls	Yes	Yes	Yes

Notes: The dependent variables are expressed in logarithmic form. tandard errors clustered at the district level in parentheses. Road refers to black topped (pucca) roads in the villages.* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.8: Treatment Effect of Bank Expansion on Employment by Distance to Population Center

	Panel A: Close to Population Center		
	(1)	(2)	(3)
	Cultivators	Agri. Laborers	Other Workers
Conventional	-0.252*	-0.379**	0.133
	(0.149)	(0.189)	(0.128)
Bias-corrected	-0.305**	-0.425**	0.170
	(0.149)	(0.189)	(0.128)
Robust	-0.305*	-0.425*	0.170
	(0.173)	(0.218)	(0.152)
Observations	246,666	221,712	243,872
State FE	Yes	Yes	Yes
Village Controls	Yes	Yes	Yes

	Panel B: Far from Population Center		
	(1)	(2)	(3)
	Cultivators	Agri. Laborers	Other Workers
Conventional	-0.217	0.109	-0.088
	(0.180)	(0.216)	(0.118)
Bias-corrected	-0.255	0.125	-0.096
	(0.180)	(0.216)	(0.118)
Robust	-0.255	0.125	-0.096
	(0.211)	(0.255)	(0.141)
Observations	234,483	192,392	226,086
State FE	Yes	Yes	Yes
Village Controls	Yes	Yes	Yes

Notes: The dependent variables are expressed in logarithmic form. Standard errors clustered at the district level in parentheses. Distance refers to the minimum distance to the nearest municipality with a population above 10k in 2011. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.9: Treatment Effect of Bank Expansion on Farmers by Crop Suitability Quartiles in 2001

	Panel A: Crop Suitability, High Input Use			
	(1) Quartile 1	(2) Quartile 2	(3) Quartile 3	(4) Quartile 4
Conventional	-0.057	-0.125	0.009	0.033
	(0.218)	(0.226)	(0.216)	(0.218)
Bias-corrected	-0.047	-0.158	0.030	0.070
	(0.218)	(0.226)	(0.216)	(0.218)
Robust	-0.047	-0.158	0.030	0.070
	(0.260)	(0.263)	(0.251)	(0.254)
Observations	193,400	125,903	115,745	79,052
State FE	Yes	Yes	Yes	Yes
	Panel B: Crop Suitability, Low Input Use			
	(1) Quartile 1	(2) Quartile 2	(3) Quartile 3	(4) Quartile 4
Conventional	-0.126	0.060	0.013	-0.070
	(0.221)	(0.225)	(0.219)	(0.244)
Bias-corrected	-0.129	0.040	0.023	-0.061
	(0.221)	(0.225)	(0.219)	(0.244)
Robust	-0.129	0.040	0.023	-0.061
	(0.260)	(0.261)	(0.261)	(0.293)
Observations	174,998	128,866	123,187	88,554
State FE	Yes	Yes	Yes	Yes

Notes: The dependent variables are expressed in logarithmic form. Standard errors clustered at the district level in parentheses. Crop suitability refers to cereal crop potential production measure (low/high input usage, log) from the FAO Global Agro-Ecological Zones (GAEZ) at the village level. Villages are then sorted into quartiles based on their crop suitability within a sub-district.* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.10: Treatment Effect of Bank Expansion on Employment by Road in 2001

Panel A: With Road

	(1) Cultivators	(2) Agri. Laborers	(3) Other Workers
Conventional	-0.023	0.221	0.062
	(0.262)	(0.261)	(0.322)
Bias-corrected	-0.046	0.275	0.076
	(0.262)	(0.261)	(0.322)
Robust	-0.046	0.275	0.076
	(0.297)	(0.313)	(0.374)
Observations	282,238	260,027	278,966
State FE	Yes	Yes	Yes

Panel B: No Road

	(1) Cultivators	(2) Agri. Laborers	(3) Other Workers
Conventional	-0.162	0.124	-0.362
	(0.244)	(0.196)	(0.235)
Bias-corrected	-0.184	0.139	-0.410*
	(0.244)	(0.196)	(0.235)
Robust	-0.184	0.139	-0.410
	(0.274)	(0.237)	(0.262)
Observations	230,867	192,298	213,160
State FE	Yes	Yes	Yes

Notes: The dependent variables are expressed in logarithmic form. Standard errors clustered at the district level in parentheses. Road refers to black topped (pucca) roads in the villages.* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.11: Treatment Effect of Bank Expansion on Employment by Distance to Population Center in 2001

	Panel A: Close to Population Center		
	(1) Cultivators	(2) Agri. Laborers	(3) Other Workers
Conventional	-0.024	0.157	0.072
	(0.295)	(0.249)	(0.336)
Bias-corrected	-0.032	0.250	0.137
	(0.295)	(0.249)	(0.336)
Robust	-0.032	0.250	0.137
	(0.342)	(0.293)	(0.389)
Observations	260,447	239,604	254,787
State FE	Yes	Yes	Yes

	Panel B: Far from Population Center		
	(1) Cultivators	(2) Agri. Laborers	(3) Other Workers
Conventional	0.025	0.371*	0.059
	(0.197)	(0.217)	(0.244)
Bias-corrected	0.035	0.414*	0.070
	(0.197)	(0.217)	(0.244)
Robust	0.035	0.414	0.070
	(0.228)	(0.255)	(0.291)
Observations	257,148	215,378	241,956
State FE	Yes	Yes	Yes

Notes: The dependent variables are expressed in logarithmic form. Standard errors clustered at the district level in parentheses. Distance refers to the minimum distance to the nearest municipality with a population above 10k in 2001. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.12: Treatment Effect of Bank Expansion on Num. of Farmers by Crop Suitability Quartiles

	Panel A: Crop Suitability, High Input Use			
	(1)	(2)	(3)	(4)
	Quartile 1	Quartile 2	Quartile 3	Quartile 4
Conventional	-37.730*	-48.448*	-15.395	0.134
	(22.141)	(25.772)	(25.825)	(25.600)
Bias-corrected	-42.863*	-57.594**	-20.345	-0.344
	(22.141)	(25.772)	(25.825)	(25.600)
Robust	-42.863*	-57.594**	-20.345	-0.344
	(24.392)	(28.747)	(27.959)	(29.500)
Observations	187,549	121,840	111,851	76,669
Mean of dependent variable	150.1	155.6	161.8	176.5
State FE	Yes	Yes	Yes	Yes
Village Controls	Yes	Yes	Yes	Yes

	Panel B: Crop Suitability, Low Input Use			
	(1)	(2)	(3)	(4)
	Quartile 1	Quartile 2	Quartile 3	Quartile 4
Conventional	-47.861*	-23.161	-19.212	7.961
	(25.690)	(25.789)	(23.696)	(27.756)
Bias-corrected	-54.596**	-30.087	-22.573	7.782
	(25.690)	(25.789)	(23.696)	(27.756)
Robust	-54.596*	-30.087	-22.573	7.782
	(28.503)	(28.105)	(26.952)	(32.252)
Observations	169,471	124,408	119,285	85,948
Mean of dependent variable	157.9	154.3	156.3	165.8
State FE	Yes	Yes	Yes	Yes
Village Controls	Yes	Yes	Yes	Yes

Notes: Standard errors clustered at the district level in parentheses. Crop suitability refers to cereal crop potential production measure (low/high input usage, log) from the FAO Global Agro-Ecological Zones (GAEZ) at the village level. Villages are then sorted into quartiles based on their crop suitability within a sub-district.* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.13: Treatment Effect of Bank Expansion on Num. of Employment by Road

	Panel A: With Road		
	(1)	(2)	(3)
	Cultivators	Agri. Laborers	Other Workers
Conventional	-31.440	-60.968	30.564***
	(28.910)	(39.333)	(10.242)
Bias-corrected	-40.391	-62.958	35.753***
	(28.910)	(39.333)	(10.242)
Robust	-40.391	-62.958	35.753***
	(32.044)	(49.980)	(10.240)
Observations	325,515	301,575	317,864
Mean of dependent variable	187.4	170.6	122.3
State FE	Yes	Yes	Yes
Village Controls	Yes	Yes	Yes

	Panel B: Without Road		
	(1)	(2)	(3)
	Cultivators	Agri. Laborers	Other Workers
Conventional	11.176	15.579	18.338
	(35.284)	(30.371)	(11.886)
Bias-corrected	15.535	13.958	17.392
	(35.284)	(30.371)	(11.886)
Robust	15.535	13.958	17.392
	(43.961)	(35.546)	(15.579)
Observations	174,778	142,253	162,925
Mean of dependent variable	104.8	69.01	50.32
State FE	Yes	Yes	Yes
Village Controls	Yes	Yes	Yes

Notes: Standard errors clustered at the district level in parentheses. Road refers to black topped (pucca) roads in the villages.* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.14: Treatment Effect of Bank Expansion on Num. of Employment by Distance to Population Center

	Panel A: Close to Population Center		
	(1)	(2)	(3)
	Cultivators	Agri. Laborers	Other Workers
Conventional	-39.229	-56.577	39.262**
	(28.147)	(42.011)	(15.702)
Bias-corrected	-48.330*	-55.803	45.032***
	(28.147)	(42.011)	(15.702)
Robust	-48.330	-55.803	45.032***
	(30.939)	(54.827)	(15.347)
Observations	253,282	233,245	247,606
Mean of dependent variable	165.4	160.8	127.8
State FE	Yes	Yes	Yes
Village Controls	Yes	Yes	Yes

	Panel B: Far from Population Center		
	(1)	(2)	(3)
	Cultivators	Agri. Laborers	Other Workers
Conventional	-1.057	-0.623	16.472*
	(21.191)	(27.381)	(9.845)
Bias-corrected	-3.183	-6.640	17.857*
	(21.191)	(27.381)	(9.845)
Robust	-3.183	-6.640	17.857
	(25.184)	(32.173)	(11.325)
Observations	247,420	210,931	233,565
Mean of dependent variable	150.4	107.8	65.51
State FE	Yes	Yes	Yes
Village Controls	Yes	Yes	Yes

Notes: Standard errors clustered at the district level in parentheses. Distance refers to the minimum distance to the nearest municipality with a population above 10k in 2011. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$