European Bank
for Reconstruction and Development

Banking and growth: evidence from a regression discontinuity analysis

Nathaniel Young

Summary
This paper investigates banking expansion and economic growth. Contrary to theoretical, cross-country and historic evidence from the United States, several recent microeconomic studies from developing country settings do not find enduring banking effects. I exploit the exogenous expansion of bank branches in India, driven by a previously unstudied policy reform from 2005. Iterating a regression discontinuity design, I trace branch growth before and after the reform along with responses from the real economy. I find strong causal evidence that the expansion of financial intermediation led to positive outcomes in both agriculture and manufacturing, and confirm growth in local GDP using nightlights data.

Keywords: banking, growth, regression discontinuity, India

JEL Classification: G21, G28, L13, O12, O16

Contact details: One Exchange Square, London EC2A 2JN, United Kingdom
Phone: +44 20 7338 8540; *Fax*: +44 20 7338 6111; *Email*: youngna@ebrd.com

Nathaniel Young is a Principal Economist at the European Bank for Reconstruction and Development.

This paper is based on my PhD dissertation at Boston University. I gratefully acknowledge financial support from the Center for Finance, Law and Policy, the Institute for Economic Development and a Boston University Department of Economics Summer Research Grant. Thank you to my advisers Dilip Mookherjee, Marc Rysman, Francesco Decarolis and Samuel Bazzi for their excellent guidance and support. I also wish to thank Kehinde Ajayi, Treb Allen, Gabriela Aparicio, Rajeev Dehejia, Jordi Jaumandreu, Jason Kerwin, Anjini Kochar, Andy Newman, Mark Rosenzweig, Johannes Schmieder and seminar participants at Aalto University, Boston University, the EBRD, EconCon, the FDIC, the FTC, IGIDR, Luxembourg School of Finance, and NEUDC for their helpful comments. Additional thanks to the scholars and researchers at ISI-Kolkata, NIPFP-Delhi, IGIDR-Mumbai, IFMR-Chennai, CDS-Thiruvananthapuram for hosting me during the early stages of this work and the many helpful conversations, and the officials at the Reserve Bank of India for their excellent help with interpreting regulations and understanding the data. All opinions expressed are those of the author alone and do not necessarily reflect the opinions of the European Bank for Reconstruction and Development. All errors are my own.

Working Paper No. 207 **Prepared in November 2017**

1 Introduction

Strong financial systems contribute to the economic health and growth of an economy (King and Levine, 1993; Jayaratne and Strahan, 1996; Rajan and Zingales, 1998). In developing economies, improving access to credit and formal savings through the expansion of the banking system is seen as an important step towards this growth and broader economic inclusion. However, the empirical findings from many financial inclusion endeavours yield a discouraging view of banking access for promoting development. In contrast, the historical expansion of bank branching in the United States, as well as recent reforms to interstate branching, provides evidence of positive effects from banking on economic growth (Dehejia and Lleras-Muney, 2007; Krishnan et al., 2014). Microfinance, once the standard bearer for financial inclusion in development contexts, is now being rethought following mixed empirical evidence on its long-term benefits (Banerjee et al., 2015). Evidence in Fulford (2013) and Kaboski and Townsend (2011, 2012) found that increased access to rural banking in India and microfinance in Thailand, respectively, increased consumption in the short run that nonetheless eroded in the medium term, raising further concerns over the potential of financial development as a vehicle for sustained growth. Moreover, several recent studies (Kochar, 2011; Panagariya, 2006; Fulford, 2013) call into question the causal identification underlying the seminal results in Burgess and Pande (2005) on the benefits of bank access in a developing country setting.

I analyse a previously unstudied policy reform in India introduced in 2005, finding new evidence that bank branch expansion had a positive effect on economic growth. The reform encouraged additional commercial bank branches to be opened in certain under-banked districts and not others, similar to policies from the Social Banking period but with important differences. The incentives created by the reform facilitate a new identification strategy applicable in the post-1990 banking environment in India. This analysis helps bridge the gap between the literature promoting financial inclusion and the frequently negative evidence from micro empirical analyses. Differences in the incentives generated by earlier reforms and those in 2005, combined with the emergence of non-government-owned (private sector) banks, explain the positive effects from this recent reform.

The responses within the banking sector that I find are concentrated in the private sector, which was largely inert before the 1990s. The expansion of the bank branch networks under the Social Banking period of the 1970s and 1980s, on which many empirical studies focus, occurred through nationalised, public sector banks and the regional rural banks they sponsor. By most accounts, these banks did not function well as financial intermediaries to mobilise savings and finance creditworthy projects. Higher incidence of government ownership in banking has been shown to correlate with slower growth looking across countries (La Porta et al., 2002). In a series of joint and separate papers, Banerjee, Cole and Duflo examine the activity of banks from the public sector in India specifically (Banerjee and Duflo, 2001; Banerjee et al., 2004; Banerjee and Duflo, 2014; Cole, 2009). They show evidence of under-lending to productive firms, inertia in credit limits extended to firms and little difference in delivering development-oriented lending resulting from government ownership. The main argument for these effects are misaligned incentives within banks.

The reform used in this study occurred in a dramatically different banking environment. Following the reforms to the banking sector in the early 1990s, a viable sector of privately owned banks emerged, while all banks enjoyed greater freedom in setting interest rates and allocating credit. This banking environment, therefore, more closely reflects India today, with private sector banks playing an increasingly important role in the economy.

My identification strategy allows me to leverage geographic and temporal variation to isolate the causal effects of expanded banking access on economic growth. I exploit the process used by regulators to select the set of under-banked districts in the 2005 reform. The selection rule, based on district population per branch relative to a statistic termed the "national average", admits a regression discontinuity design. The national average constitutes a threshold, where districts with higher populations per branch receive treatment and the others do not. Incentives set in place by the reform encouraged additional branch entry in treated districts. The regression discontinuity allows me to overcome the classic endogeneity concern of bank branch expansion selecting on growth potential, and to separate out effects from other interventions. Despite evolving values of the national average and branches per capita in later years, the official list of under-banked districts remained essentially unchanged over the course of the policy. I am therefore able to trace the policy effect on a variety of outcomes through time, with pre-reform years serving as placebos. I estimate the average treatment effect of the reform through time by separately estimating the regression discontinuity for each year from 2002 to 2012.

I draw on several different sources for the data in this analysis, including India's central bank, the Reserve Bank of India (RBI), the Ministry of Agriculture, India's Annual Survey of Industries (ASI) and remote sensing data on the amount of light emitted at night and measures of rainfall. The detailed data from the RBI on bank branches and credit, from separate datasets, help provide a cross-check for my two broad banking outcomes. I combine separately reported data on district-level crop production statistics and farm harvest prices from the Ministry of Agriculture to examine responses in agricultural outcomes. Together, the datasets allow analysis into India's banking sector and two major productive sectors.

My first set of results verify that the policy reform resulted in a significant expansion of bank branches by the private sector in under-served areas. The cumulative effect of the reform is estimated as an average additional 10 private bank branches per district by the start of 2012. The effect is large, approximately 50 per cent of the sample average of operating private branches per district in districts around the threshold. I consider the effects of this entry on competition, the strategic responses of banks and implications for the delivery of credit. I examine these effects using the panel nature of my branch and credit data.

The time frame outlined in the policy reform governing branch expansion generates separate effects from credit lines and bank branching. Banks were able to delay additional branch openings in under-banked districts for a specific period due to gaps between the policy implementation date, the last date to submit expansion plans and the length of branch licence validity. Private bank branching in under-banked districts remained low during that window, climbing steadily afterwards.

Meanwhile, private credit in the affected districts expanded immediately with the policy implementation. The pattern of these responses is consistent with anticipation and competition effects between banks in contested districts, highlighting private sector banks as strategic players following profit-maximising objectives. I provide a theoretical framework of adverse selection and switching costs to outline how the anticipation of intensified competition led banks to expand credit and lock in consumers who face positive switching costs. In 2006, the reform had already induced an average increase of 6,220 private sector credit accounts for under-banked districts, approximately 48 per cent of the sample average around the threshold. No response from public sector lending is observed.

If banks contribute to economic growth, then we would expect to observe credit flows to the productive sector. The null, if not negative, medium-term effect of expanded rural bank access from Fulford (2013) revolved around intertemporal consumption smoothing. Loans made to productive uses should avoid this pitfall, however, if they improve productive processes.[1] Fortunately, the disaggregation of credit data reported by the RBI into economic sector and population group allows me to examine lending to agriculture and manufacturing separately.

Importantly, I find positive growth in the credit extended for agricultural use in rural and semi-urban areas of under-banked districts near the cutoff. Agriculture constitutes a major employment activity in India, with over 56 per cent of workers in 2001 engaged in agricultural endeavours. Further, policy-makers placed particular importance on the availability of credit to rural and semi-urban agriculture leading up to the reform. These results demonstrate that reform effects were not solely concentrated in high population areas, contributing, at least partially, to financial inclusion in rural areas.

Next, I estimate the effects of the reform on agricultural outcomes, finding positive effects consistent with the expansion of credit. Positive responses in yield (output per hectare) and raw output are estimated for several important crops including rice, wheat, cotton and onions. I construct a revenue weighted index of crop yields by combining crop statistics with district level harvest prices. This index helps account for differences in the importance of individual crops across districts. I estimate that an increase of 1,000 private bank credit accounts in a district raises average crop yield by 2.3 per cent. This effect is a little less than one-third of the effect from a positive rainfall shock on yield found in Jayachandran (2006).

Banking effects in farming have been found in the farm labour supply (Jayachandran, 2006), and in cropping decisions in recent work by Allen and Atkin (2015). The improved yields supply evidence that banks are leading to actual improvements in agricultural outcomes. These effects may be accruing through either the adoption of higher quality inputs purchased with credit, such as fertilizers and machinery, or solely through the shuffling of crops across land with differing crop-specific yield potentials.

The other key productive sector in India that I examine is manufacturing, using annual data from the ASI. I find that enterprises in states with populations most affected by the reform experienced faster growth in their use of resources. Specifically, those enterprises reported higher total investments, working capital and capital labour ratios. The survey response on outstanding loans

[1] Personal loans may improve productivity as well, if they contribute to building human capital, such as through expenditures on education or health.

also showed these enterprises had 23 per cent higher loans than the control group. This finding supports estimates on increased credit access to manufacturing using the RBI data, and suggests that the heightened investments occurred through financing. Since the ASI data that I analyse are available at the state level, I must follow a different empirical strategy for this analysis. Thus, results should be interpreted with some caution.

The results from manufacturing are largely consistent with two analyses examining effects of branching in the United States. Dehejia and Lleras-Muney (2007) find that increased bank branching in the United States from 1900-40 encouraged growth in agricultural and manufacturing. Krishnan et al. (2014) show that increased branching activity in the United States, following the Interstate Banking and Branching Act of 1994, led to greater efficiency gains by previously credit constrained manufacturers. The increases to investments and capital intensity found in my analysis implies greater access to capital in manufacturing following the expansion of credit. These results are related to resource misallocation and lower aggregate total factor productivity (TFP) examined by Hsieh and Klenow (2009), implying potential efficiency gains from the reform through reallocations of capital.

Lastly, I confirm the aggregate effect on local GDP growth by showing that areas with expanding banking services experienced higher rates of growth in nighttime light intensity in the years following the reform. The nightlights data provide a reliable proxy for economic growth to overcome the lack of regularly available data on district level GDP in India. Taking the elasticity of nighttime light to GDP estimated in Henderson et al. (2012), I estimate that each additional private bank branch led to a 0.33 per cent increase in local GDP. Overall, these findings offer strong causal evidence that the expansion of the financial system facilitates growth across productive sectors and encourages economic development.

In the next section I describe the important institutional aspects of India's banking system and the policy reforms to the branch licensing policies utilised for analysis. Section 3 outlines a simple theoretical framework of bank responses to the reform, section 4 discusses the empirical strategy and section 5 describes the data used in analysis. In section 6 I present the empirical results and section 7 concludes.

2 Policy reform and institutional background

2.1 Policy reform

The analysis in this paper utilises a policy reform to bank branch licensing in India implemented on 8 September 2005. The banking sector in India does not permit free entry of banking firms or branches. New bank licences are granted infrequently by the Reserve Bank of India (RBI), India's central bank, through special campaigns with recent waves in the early 1990s and again in the early 2000s. Banks must also acquire licences prior to opening all new branches, as well as receive permission to close or shift branches in most markets. Prior to the 2005 reform, banks applied for each of these changes on a case-by-case basis through the regional office of the RBI. No broad directive with regards to the composition of markets served by the bank, such as a requirement to open branches in rural areas, existed following the end of the Social Banking period in 1990.[2]

The reform in 2005 changed the regulatory environment in two fundamental ways. First, the total branch licences issued to a bank was tied to their proposed entry in a set of districts the RBI designated as being under-banked.[3] The rule governing the assignment of under-banked status was based on the district average persons per branch relative to the national population per branch for India (RBI, 2009). Though not stated explicitly, I will argue that a form of quota system operated that required expansion in under-banked districts for entry into rich markets. The randomisation of districts around the national average cutoff, resulting in extra branching incentives for those falling on the under-banked side, provides the identifying variation exploited in this analysis. I discuss this strategy in detail in section 4. Second, the case-by-case application procedure followed by banks to request new licences was replaced with an Annual Branch Expansion Plan (ABEP) approach. Banks proposed a set of branch openings, closings and shifts to be implemented over the next year. The RBI reviewed the list centrally, potentially meeting with bank management, and granted the set of permissions (Master Circular, 2005).[4]

Important differences exist between the above policy and those implemented under Social Banking. Banks experienced far greater choice in selecting locations in which to open under the 2005 reform compared with the Social Banking rules. Unlike the 4:1 entitlement policy studied in Burgess and Pande (2005), which required intervention branches to be opened strictly in unbanked markets, banks could choose among any markets within under-banked districts to satisfy their obligation. That characteristic created the potential for increased direct competition between banks. In stark contrast to the planned approach to district-wise branch expansion implemented in the 1980s (RBI, 2009; Kochar, 2011), banks under the current reform chose which under-banked

[2]The LEAD banking scheme was in operation during this time, however, by which one bank was assigned to each development block and made responsible for meeting agreed levels of branching and banking services. These banks were typically selected from the set of government owned banks. The service area approach (SAA) also operated at this time, partitioning rural areas between banks for implementing development objectives.

[3]Banks were also judged on their provision of "no-frills" accounts, meeting priority sector lending obligations and their handling of complaints (Master Circular, 2005).

[4]Permissions were valid for one year with the potential for extensions. Banks accomplishing 75 per cent of their planned expansions could submit their next ABEP regardless of the lapsed time.

districts to enter, as well as decided by how much to expand their total branch network.

Lastly, the banking environment differed drastically in its composition and scope of business. The private sector, largely inert during Social Banking, expanded and gained vitality following the deregulations beginning in 1990 and the infusion of "new private" banks. Government-owned banks, consisting of the State Bank of India and its Associated Banks, the set of nationalised banks, and most regional rural banks (RRBs), traditionally dominated the banking system in India. In recent years, private sector banks have been operating alongside and competing with government-owned banks. The new private banks broadly face the same regulation as the other scheduled commercial banks.[5] The other policies they face, as well as their requirements to the Priority Sector lending scheme, are identical to those for the SBI and Nationalised banks. RRBs and foreign banks face tailored regulations, including those pertaining to branching requirements.

2.2 Policy details and timing

The list of under-banked districts remained nearly constant through the end of the sample period, with minor revisions to the 2005 list issued in 2006. The list was then reissued unchanged from 2007 to 2010.[6] After 2010, certain states were made ineligible for under-banked status, reducing the number of under-banked districts, but not introducing any new districts to under-banked status. Although additional reforms altered the incentives for branch expansion both within and outside under-banked districts, given the lagged nature of branch openings to licence issuance, I find lasting effects through 2012 as expected. The persistence of under-banked status helps in identifying the policy effect on banking, and in turn the effects of banking on real economic outcomes. I discuss how the empirical methodology is designed to exploit this fact in section 4.

Further, heavy regulations on the closing and shifting of branches limited subsequent adjustments of existing branches in a bank's network.[7] Few branch closures are observed in the data. During bank mergers, most branches of the exiting bank are reopened under the acquiring bank, though some branches do get converted to satellite offices and others are permanently closed. Limited branch exit discourages strategies to meet regulation requirements that would entail moving temporary branches across districts each year.

The implementation of the reform, as well as the process of its drafting, created opportunities for banks to behave strategically in timing their responses. In the online appendix I discuss grace

[5] Private sector banks carry the additional mandate of maintaining at least 25 per cent of their branch network in population centres with fewer than 100,000 people.

[6] Starting in 2008, certain centres within under-banked districts were made ineligible to count towards a bank's serving of common persons. Specifically, centres within the municipal limits of state capitals, district headquarters and metropolitan centres were deemed ineligible. Further, centres within 100 km of Mumbai, New Delhi, Kolkata and Chennai, and 50 km of state capitals were ineligible. Exceptions were made for the state of Jammu and Kashmir, and the seven north-eastern states, Arunachal Pradesh, Assam, Manipur, Meghalaya, Mizoram, Nagaland and Tripura.

[7] Branches were not allowed to shift outside otherwise unbanked centres. Given that a location was served by another commercial bank branch (other than an RRB), a branch could only shift to centres in the same or lower population group classification, and in the case of branches in under-banked districts, could only shift to centres within under-banked districts.

periods, licence validity length, the Vyas Committee and Service Area Approach reform.

2.3 Policy reform discussion

The 2005 branch licensing policy reform purposefully created new incentives for scheduled commercial banks to open in centres conditional on their district's under-banked status. Licences for branches in high profit potential centres in banked districts were used to leverage bank entry into under-banked districts. This mechanism works most effectively during periods of high demand for bank branches in "rich" areas, as was presumably the case experienced in India during its time of high economic growth beginning in 2003 and continuing through the decade.

The branching policies and reform placed no requirements on the amount of banking required to occur at each branch. There are staffing requirements for branches, as well as minimal days and hours of operation. Banks must also offer "no-frills" accounts that carry limited fees and low minimal balances to prevent the exclusion of poor customers. Despite these requirements, though, banks could maintain staffed branches that simply minimised costs by not reviewing or approving any loan applications, not pursue new customers, and only accept deposits.

An important regulation affecting bank lending behaviour is the set of Priority Sector lending ratios.[8] These requirements must only be met at the bank level, however, such that some branches may carry heavy amounts of priority sector loans while others lend nothing at all. In 2007, new guidelines were adopted for the priority sector, reducing the set of loan categories eligible for priority status.[9] The reformed guidelines concentrated lending into direct and indirect agricultural endeavours, and limited the amount going to microfinance institutions and other modes of on-lending. While the adjustments to the priority sector requirements still applied at the bank level, and not by geography, I will consider potential effects from this reform in an analysis of loans by category.

[8] Banks must maintain 40 per cent of their outstanding credit in loans to the priority sector. Banks failing to meet their 40 per cent requirement must make up the difference with loans to the NABARD RIDF fund at deterrent rates. Banks typically come very close to meeting the requirement, overshooting slightly in some years and falling short in others.

[9] An earlier set of reforms to the composition of the priority sector occurred in 1998 and 2000, studied in Banerjee and Duflo (2014).

3 Theoretical framework

This section articulates a simple theoretical framework to provide intuition for the effects of the reform on branch entry and responses in credit levels. The theoretical framework demonstrates how the 2005 policy reform could incentivise higher rates of entry in under-banked districts and increase lending without addressing the underlying profitability conditions of those districts. Introducing switching costs on borrowers to establish credit relationships with new banks, I show how in a two-period framework the increased threat of entry may induce an expansion of credit prior to realised entry. Then considering heterogeneous entry costs for bank-district pairs, I argue the reform would lead to an expansion of branching in treated districts as banks cross-subsidised required entry in lower performing districts with entry in richer ones. Lastly, as the above mechanisms rely on incentives consistent with profit-maximising objectives, different responses to the reform by private and public sector banks are predicted.

The framework adopts a standard characterisation of financial intermediation with adverse selection of borrowers, a feature common to credit markets in developing economies.[10] Consider a single market with two periods and two types of borrowers, safe and risky. In the first period, a policy reform that will encourage entry in a (potentially unknown) set of markets beginning in the second period is announced. In the second period the reform is in effect.

As in Stiglitz and Weiss (1981), each borrower has a potential project that requires a loan (normalised to size one for all borrowers) and yields the same expected return across borrowers. The borrower is assumed to have the same potential project in each period. Assume that the return from a failed project is zero, and that $P_s(R_s^A)R_s^A = P_r(R_r^A)R_r^A$, where R_i^A is the return from a successful (denoted A) project for type $i \in \{safe\,(s), risky\,(r)\}$ and $P_i(R_i^A)$ is the probability of success for type i. Thus, safe types have projects with lower returns conditional on success but succeed with greater probability $P_s(R_s^A) > P_r(R_r^A)$.

If banks operate in the market, they can offer a standard debt contract with fixed repayment. Only loans where $R_i^A > (1+r_i) > 0$ face positive demand, and assume that borrowers face limited liability. When a project is successful the borrower pays back the principal on the loan plus interest at rate r_i, but in case of failure no payment is made and both borrower and bank receive zero. Borrowers face an outside option that provides utility equal to μ. Both borrowers and banks discount the future at rate δ and are risk neutral. While borrowers know their own type, banks only know the distribution of types and the parameters defining the projects. Banks prefer to lend to the safe types due to limited liability, but cannot distinguish between types in the general framework. Depending on the set of parameters and the share of safe and risky types in the population, banks may choose to ration credit in response to adverse selection, or the market may collapse entirely (Stiglitz and Weiss, 1981).

To capture the dynamic effect of the policy reform, consider the two following modifications: (i) banks possess a screening technology that reveals a potential borrower's type with certainty and costs amount s; and (ii) there exists a downward sloping demand curve among safe types.

[10] See Conning and Udry (2007) for a survey of approaches.

[11] The cost of screening, which banks pass on to borrowers, introduces a switching cost. In a practical sense, these costs may include the submission and review of a loan application, and efforts taken to establish a good relationship between a borrower and branch manager.[12] The downward sloping demand curve is necessary for competition to affect the size of the market served, and not just the division of market shares, since borrowers are otherwise homogenous within types.

Empirical evidence of switching costs in bank lending from Barone et al. (2011), showed that medium to large borrowing firms in Italy required sizeable premiums on interest rates to switch their main lenders in local business credit markets. Further, banks appeared to actively provide discounts to attract switching firms. Their findings are consistent with the theoretical results of the 2-bank, 2-period model in Gehrig and Stenbacka (2007), where banks compete for borrowers with individual-specific switching costs. The current framework is similar to the Gehrig and Stenbacka (2007) model, which also incorporated adverse selection, with the important difference that here switching costs are assumed to be constant across borrowers, banks cannot price discriminate between new and old borrowers, one bank may be an incumbent and a costly screening mechanism replaces learning borrower types during the first period of lending.[13] These assumptions will be appropriate if loan officers have less liberty to adjust interest rates from those set at the bank level for small loans, which seems plausible for the context. This framework abstracts from the churn of customers between banks, since bank-borrower and loan-level data are unavailable, focusing instead on dynamic effects in total credit amounts for markets with switching costs and anticipated entry.

To simplify the analysis, assume parameters are such that banks always choose to screen borrowers and never find it profitable to lend to the risky types.[14] Adding the assumptions that borrowers must repay the full amount of the loan conditional on a successful project, and that borrowers cannot accept contracts with the potential for negative consumption in any period, the expected default rate from safe types will be straightforward and banks will know the demand conditional on the interest rate offered with certainty.[15] This assumption greatly simplifies the game as it allows the borrower's decision process to be considered separately for each period, since agents cannot accept negative first period expected returns to gain access to more favourable expected lending conditions offered in the future.

Assume banks are symmetric and profit maximisers, each facing an exogenous marginal cost of funds, including administrative costs from lending. Recall that banks cannot discriminate in the interest rate they offer to repeat versus first time borrowers. Since banks observe the parameters

[11] A wide range of assumptions can satisfy this condition, for example, if personal costs of marketing the successful project differs between borrowers then demand for loans will be non-increasing in r_s.

[12] Klemperer (1987) mentions banks as a motivating example in his seminal work on switching costs. A survey on switching costs may be found in Farrell and Klemperer (2007).

[13] Paying for screening could be viewed as replacing the costs of lower returns from serving risky borrowers due to adverse selection before banks learn their types.

[14] Vesala (2007) presents a model of adverse selection and switching costs where relationship lending leads to a noisy signal on borrower quality, with banks optimally choosing to accept fractions of applicants with either signal, a consideration beyond the scope of this analysis.

[15] A contract with potential negative consumption would arise when limited liability protects the borrower against a failed project, but not from a successful project for a borrower whose high marketing costs leaves them less from the project than the fixed payment owed to the bank.

on the population defining the distribution of safe types, they know the slope of the demand curve, though they do not know any particular borrower's value of the loan. Without the threat of entry, a monopolist serving the market in the first period maximises profits by serving the same set of borrowers in each period, increasing the interest rate in the second period to capture the additional surplus the borrowers receive from not paying the screening cost again (a sketch of the proof is given in the online appendix to this paper). Knowing this, the monopolist may work backwards from the second period to determine the profit-maximising interest rates in each period. In contrast, when two banks serve a market, they compete in prices. If both enter the market in the same period, then each offers the zero profit interest rate and they split the market.

However, if one bank acts as an incumbent, then it may choose to alter its behaviour when anticipating the potential of entry. The screening cost operates as a switching cost for the borrower as previously discussed. Borrowers will go to whichever bank results in them keeping the highest expected return from their project. For first time borrowers this is simply the bank offering the lowest interest rate. Repeat borrowers must compare their expected payoff from the incumbent's second period interest rate to that of the entrant plus the screening fee required to switch. The resulting equilibrium is intuitive: in the second period, under-cutting leads the entrant to offer the zero profit interest rate and the incumbent offers an interest rate making its set of first period borrowers indifferent between switching to the entrant and staying. Since the set of first period borrowers is entirely determined by the first period interest rate, the second period interest rate is a function of the first period interest rate and the screening cost. Knowing this, the incumbent chooses the first period interest rate that maximises profits over both periods. The threat of entry will result in the monopolist offering lower first period interest rates to secure a larger base of customers from which to earn positive profits in the second period. The set of parameters will determine how willing the incumbent is to trade off first period profits for those in the second period. The entrant will serve the remainder of the market that demands loans at the zero profit condition. Thus, credit will initially expand with the announcement of the policy reform and again upon realised entry.[16]

3.1 Entry

The effects on entry must be primarily driven through changes to the structure of fixed costs of entry as the reform did not otherwise target local market conditions. Consider multiple markets described by the framework above. Markets are differentiated by their set of parameters already discussed plus overall market size. Suppose banks each draw market-specific fixed costs of entry for every market. Abstracting from the strategic considerations of entry, assume banks act myopically such that they expect to act as a monopolist if entering a market unbanked in the first period or as a duopolist when entering banked markets. Under these assumptions, expected profits for each market is known to a bank and entry will occur for all markets j satisfying $E[\pi_B^j] - F_j > 0$, where F_j denotes the fixed cost in market j. Markets with low profit potential or high fixed entry costs will fail to attract banks.

Consider a rule that ties permission for entry in some high profit potential markets to entry in

[16]Additional discussion of the theoretical framework is available in the online appendix to this paper.

lower profit ones. Banks facing binding constraints will now open into markets where $E[\pi_B^{UB1}] - F_{UB1} < 0$ if these losses may be offset by the profit gains from the rich market, $E[\pi_B^j] + E[\pi_B^{UB1}] - F_j - F_{UB1} > 0$. This condition will be more easily satisfied in policy eligible districts with higher expected profits that faced high fixed entry costs. Once entered, however, these markets may produce high levels of banking activity. In contrast, the set of markets originally served without the reform may contract if the lowest profit earning locations cannot offset the losses from policy eligible markets. Lastly, the joint positive profits will be hardest to satisfy for policy eligible districts that face the lowest profit potential and highest fixed costs of entry. The reform will be unlikely to produce positive banking results for such markets. Note, the above implies entry may be most profitable in locations where banks open as a competitor, with lower fixed costs making up for stronger competition for borrowers. Thus, both entry as monopolists and as competitors is possible.

To the extent that population per branch, upon which the 2005 policy reform is based, provides a suitable proxy for potential profitability of a district, responses that should hold true for local averages in branch entry along this measure may be predicted. Districts in the lower tail of population per branch (the most heavily banked districts) will likely continue to experience branch growth.[17] Districts just below the cutoff should not experience higher growth rates than in the policy's absence. The incentive to open into these districts is diminished as they offer, on average, the lowest profit potential of banked districts and are therefore relatively costly as they would still count against a bank's quota of openings. In contrast, districts just above the cutoff, such that they receive treatment status, are likely to be the most profitable on average. Moving down the tail of population per branch will represent districts with lower and lower profit potential, making them unlikely to experience a benefit from the reform.

3.2 Predictions

The above framework suggests three main empirically testable predictions of banking responses to the policy reform.

Prediction 1. Branch entry will increase the most from the reform in under-banked districts just above the cutoff. Entry is less likely to occur just below the cutoff for untreated districts, and the least likely to occur in the tail of under-banked districts, despite treatment status. Changes to branch entry in the tail of heavily banked districts is ambiguous. Growth will be likely, however, as growth in these districts fuels the responses elsewhere.

Prediction 2. The amount of credit will expand in districts where increased entry under the reform is expected to occur. Credit will initially expand with the announcement of the policy

[17] Branching may occur at higher or lower rates than in the years leading up to the reform, conditional on overall economic growth. If the regulatory body relaxes requirements in these markets to encourage additional entry in under-banked districts, then branching rates will increase. If the requirement to open in under-banked areas is set too strictly, then rates in these high profit districts may decrease.

reform, and again upon realised entry. From prediction 1, this means districts just above the cutoff for treatment should experience an expansion of credit at the time the policy is revealed. The districts just below the cutoff should not experience additional expansion and may in fact stagnate. Districts in the tail of the under-banked set are unlikely to experience credit expansion due to the reform.

Prediction 3. The expansion of branching and credit should not be observable for public sector banks. This follows from the driving assumption of profit maximisation in the theoretical framework. Banks following other objective functions, as public sector banks might, would be less likely to generate the above responses.

4 Empirical methodology

Identifying the effect of bank branching on banking and real economic outcomes can be frustrated by classic endogeneity concerns outlined in previous work (Burgess and Pande, 2005). The unique policy aspects of the 2005 branching reform create an environment facilitating the identification of banking effects on agricultural, industrial and other outcomes. I am able to separately identify the banking effects from simultaneously operating reforms and confounding factors by employing a regression discontinuity design. Further, the assumptions supporting the RD analysis are at least partially testable.

4.1 Regression discontinuity

The method employed by the RBI for identifying districts as under-banked in the 2005 branching policy reform, based on simple district and national averages of population per branch, yields a clear quasi-natural experiment exploitable by regression discontinuity techniques. Under-banked districts were identified using two inputs. First, the national population of India, taken from the population census conducted in 2001, was divided by the total number of scheduled commercial bank branches operating in the country in 2005-06 to obtain a "national average of population per branch". Then an analogous value was calculated for each district and compared with this national average. Those districts with a calculated value higher than the national value were designated under-banked. Figure 1 shows district under-banked status from the 2006 list of under-banked districts plotted against district population per branch around the national average. According to the rule, districts to the right of the cutoff should be assigned under-banked status, as is broadly confirmed in the graph.[18] A map of the districts in India with their corresponding district averages is presented in the left map in panel A of Figure 2.[19]

The above algorithm induces a cutoff at the value of the national average, treating district population per branch as the forcing variable. The policy generates an arbitrary difference in districts falling on the "under-banked" side of the cutoff, which offers an additional value to banks opening branches within their borders: such openings count towards their requirement for "serving common persons" in order to gain permissions for branches in rich markets. Districts falling on the other side of the threshold do not offer this benefit, despite being similar along other dimensions. Thus, the policy effects the probability that the districts will receive additional branches through its leveraging of bank incentives. This estimation strategy will be valid if the distribution of potential outcomes is continuous at the cutoff (Lee, 2008). I will verify two conditions in support of this validity below. First, I will check for a lack of perfect manipulation of the running variable so that agents cannot determine a district's treatment status. Then I will test whether other factors that may affect the outcomes of interest are continuous in the district population per branch near the cutoff.

[18] Five districts do not follow the assignment rule, with four of them remaining in the sample used in estimation (see the section on constructing the forcing variable in the Data Appendix for details).

[19] The districts with greater deficits of branches per person, denoted by darker colours, matches closely with the areas identified as being more broadly under served by the map from the Vyas Committee issued in 2004.

Figure 3 presents visual results from the McCrary test for manipulation of the running variable around the threshold (McCrary, 2008). The distribution of districts along the running variable is shown to be smooth around the threshold. The discontinuity estimate in the log difference in height is 6.6 with a standard error of 22, thus the test fails to reject the null hypothesis of continuity. The figure also highlights another ideal trait of this environment; the cutoff is located near the peak of the density, meaning most districts fall close to the cutoff. The even distribution of districts around the cutoff holds within regions as well, shown in panel B of Figure 2. While the central, eastern and north-eastern regions of India are relatively less banked than the north, south and west, each region has districts falling near the cutoff on both sides.

The lack of manipulation around the cutoff, beyond passing the McCrary test, is extremely defensible on intuitive grounds. Even if banks and districts were able to perfectly anticipate the criteria for assigning under-banked status, their ability to manipulate the assignment would be limited. The population level in the current equation was taken in 2001, four years prior to the policy. Thus, agents attempting to influence district status could only do so through altering the number of operating branches within district boundaries, which results from the collective branching decisions of all banks and conditional on RBI permissions, making manipulability extremely unlikely.

Figure 4 presents a series of plots of district baseline characteristics, with dots reporting local averages for districts falling within 200 persons per branch non-overlapping bins. A local linear regression of the data is shown with flexible slope on either side of the cutoff. Each of these characteristics appears to be smooth at the cutoff, suggesting proper randomization of districts around the cutoff. The continuity is tested formally by performing RD analysis with the baseline characteristics as the dependent variable. The tests fail to reject the null hypothesis of continuity at the threshold, with reduced form results presented in Table 1.

While the figures constitute a visual RD testing for continuity at the cutoff centred at zero, they also summarise broader trends in branching at the time of the policy reform. Districts left of the cutoff enjoyed more branches per person by definition. These districts also tended to be places with more highly concentrated populations, exhibited higher literacy rates, had lower populations of scheduled caste and tribe persons and had a lower percentage of main workers engaged in agriculture.[20]

4.1.1 *Technical details of RD*

The identification of local average treatment effects through regression discontinuity analysis is now well established in the literature (Black, 1999; Angrist and Lavy, 1999; Van der Klaauw, 2002; Lee et al., 2004), with the theoretical work on identification in Hahn et al. (2001) and the origins of the method in Thistlethwaite and Campbell (1960). To reduce bias from including observations far away from the cutoff where the identification does not hold, I use local linear regressions, dropping observations outside a set bandwidth of the cutoff (Hahn et al., 2001; Lee

[20] District rainfall, an important agricultural input, is also shown to be smooth at the cutoff for the years in analysis, presented in the online appendix.

and Lemieux, 2010). I restrict all analysis to local linear and local second degree polynomial regressions as recommended in Gelman and Imbens (2014). I set the bandwidth at 3,500 persons per branch for all regressions, which falls within the range of optimal bandwidths selected for individual years by the Imbens and Kalyanaraman (2011) method.[21] I fix the bandwidth to provide transparency for tracing the evolution of the policy effect across years, as this fixes the set of included districts across regressions. The map on the right in Figure 2 indicates districts included in the local linear regressions by treatment status. The treated districts are geographically distributed across most of the country and generally well mixed with control districts. A map of the districts by treatment status used in the local linear regressions is shown in the right panel of Figure 2.

For each year, I first estimate the local linear regression of the reduced form equation,

$$y_i \quad = \alpha + D_i\tau + f(PopPerBranch - Cutoff) + \delta X_i + \varepsilon_i \tag{1}$$

using a uniform kernel. y_i denotes a banking or economic outcome of interest in district i, such as the number of operating bank branches or crop yield. $D_i = 1[PopPerBranch_i - Cutoff \geq 0]$ is an indicator for satisfying the rule for assignment to under-banked status, $PopPerBranch_i$ is the population per branch for district i, $f(\cdot)$ is a flexible functional form, X_i is a set of controls, τ is the coefficient of interest measuring the discontinuity at the threshold, and ε_i is an idiosyncratic error.

In all regressions, I include the pre-randomisation assignment value of the dependent variable from 2001 in the set of controls to improve precision and reduce sampling variability (Imbens and Lemieux, 2008; Lee and Lemieux, 2010). In addition, I include the 2001 district population and its square, as well as the percentage of workers engaged in agriculture. Rainfall is an important agricultural input in much of the country and likely to affect the credit and agricultural values which may adjust quickly to realised conditions. Therefore, I include the yearly deviation of monsoon rainfall from its district mean, and the lag of this measure. The rainfall variables are excluded from the estimates on bank and branching entry as these are less likely to respond to transient shocks. The described method constitutes the reduced form estimate from a fuzzy RD design estimated via two-stage least squares, with the probability of under-banked status instrumenting for actual assignment. The estimated discontinuities are reported graphically.

I report the fuzzy RD results implementing the regression discontinuity using Calonico, Cattaneo and Titiunik's "rdrobust" package with a triangular kernel. I use the fuzzy RD because the rule assigning under-banked status does not perfectly match the realised list.[22] The triangular kernel places greater weight on observations within the bandwidth that are closer to the cutoff where districts should be most comparable. To implement the fuzzy RD analysis I first "residualise" the data, regressing y_i on the set of controls X_i from equation 1, then estimating equation 1 replacing the left-hand variable with the residuals obtained from the first regression and dropping

[21]Results are robust to different bandwidth selections, and second degree polynomials typically perform better with wider bandwidths than linear specifications as in the example from Lee and Lemieux (2010).

[22]I fail to match 5 out of 572 districts to their realised under-banked status from the 2006 list. See the data appendix for details.

the controls from the specification (Lee and Lemieux, 2010). Conventional estimates of the RD are reported, as are bias-corrected estimates and the robust standard errors from Calonico et al. (2014). I will focus on the conventional estimates and standard errors in discussing results.

4.1.2 *Instrumenting*

The relationship of greatest interest in this context exists between the economic outcomes and the realised banking environments, rather than assignment to under-banked status. For agricultural and income growth outcomes, in addition to presenting the reduced form effects of under-banked status, I estimate the effects with the fuzzy RD instrumenting for banking outcomes. That estimate will inform the effect of the specific banking outcome on the economic outcome of interest. However, that effect should be interpreted with care as the reform status will influence multiple dimensions of bank markets at once. Choosing the number of branches or credit accounts assigns the full effect of increased financial access to that one outcome. Still, any individual outcome may be taken as a proxy for the intensity of the reform treatment.

4.1.3 *Dynamic strategy*

The identification of the policy effect on banking outcomes is bolstered by the ability to regularly estimate the effect of the reform through time, both before and following its implementation. In the pre-reform period, no discontinuity should exist at the cutoff. In the post-reform period, the effect of the policy should be expected to grow according to the timing set in place by the rules of the reform and its revelation. To demonstrate the timing of the reform effects, I estimate equation 1 separately by year for banking outcomes, agricultural outcomes and measures of local economic growth from remote sensing, that is night-time light emitted into space. The list of under-banked districts from the RBI remained essentially unchanged in the reform period.[23] To reflect this, I keep the forcing variable fixed across regressions. The evolving estimates of the discontinuity therefore capture the short and medium-term policy effects as they emerge.

4.2 Manufacturing

To examine the effect of increased financial access on the manufacturing sector, I use ASI data available at the state level. The level of aggregation prevents the conducting of the regression discontinuity just described. Instead, I follow a difference in differences approach, utilising the institutional knowledge of the reforms to construct sets of treatment and control states.

I select the set of "under-banked treatment states" in the following way. Using population census data at the district level, I construct the shares of state population in under-banked districts.

[23] As noted earlier, this appears to be an administrative feature of the reform. The population per branch of the districts, as well as the national average, continued changing throughout the reform.

For the population of each state in under-banked districts, I calculate the share of that population belonging to districts falling within a close bandwidth of the national average of population per branch, generally within 4,000 persons per branch. Those states with large shares of their population in under-banked districts close to the threshold are selected as the treatment group. I then construct a control group using a comparable procedure from districts with banked status. "Banked States" include Haryana, Uttarakhand, Punjab, Mizoram, Daman and Dimiu, Karnataka, Puducherry, and "under-banked States" include Rajasthan, Tripura, Jharkhand, Orissa, Dadra and Nagar Haveli.

For each treatment and control group pairing, I estimate the following,

$$\begin{aligned} y_{it} = \quad & \alpha + \xi post06_t * treat_i + \varphi post06_t + \psi treat_i \\ & + \beta_1 year_t * state_i + \beta_2 year_t + \beta_3 state_i + \beta_4 X_{it} + \omega_{it} \end{aligned} \tag{2}$$

where $post06_t$ indicates financial years 2006 and later, $treat_i$ indicates that the state belongs to the treatment group, and the remaining terms indicate controls for state fixed effects and state specific time trends, as well as a matrix of additional controls in X_{it} with an idiosyncratic error ω_{it}. The coefficient of interest will be on the interaction term $post06_t * treat_i$, which will give the difference of within-state differences between the states receiving under-banked status and those not. In addition to controlling for post-2006 and treated state individual effects, the regressions include the logged number of manufacturing units in the firm and the logged number of employees in the enterprise to control for enterprise size. Plant age and its square are also included as controls as these may influence the firms' access to credit and capital markets. Although this identification strategy is not as ideal as the RD, the careful selection of the treatment and control states should help in eliminating potential threats and I will take the estimate as suggestive of the effect from the policy reform on manufacturing.

5 Data

The primary data on banking are from datasets maintained by the RBI. The Master Office File (MOF) provides a detailed record of bank branch locations and characteristics. The number of branches operating in each district per year are calculated from branch opening and closing dates, which are then paired with population census data to construct the running variable as well as the cutoff, the inverse of the national branches per capita. The Basic Statistical Returns 1 and 2 provide time series data on credit and deposits at various levels of aggregation.

The empirical methods and analysis pursued in this work are greatly determined by the level of data availability. Although detailed branch location data may be constructed at the daily level by bank, much of the credit and deposits data are only available annually as aggregates at the bank group level by district. Thus, matching credit data to any particular bank or branch in a district, other than cases where a single bank from a bank group serves a district, is impossible. Fortunately, the policy reform applied at the district level, allowing analysis directly at the level of the reform. Utilising the time dimension further helps to disentangle some effects of the reform from changes to bank group classifications.

The credit limits, amounts and accounts data reported to the RBI are delineated by their intended geographic area of utilisation. The use of call reports from banks do not typically allow for this level of geographic precision in terms of the utilisation of funds, distinguishing this analysis from other work. This feature also increases confidence that we are measuring the local availability of credit, as loans are less likely to be financing projects in neighbouring districts.

To conduct the analysis on agriculture, I develop a new dataset by processing and combining separate annually available data from the Ministry of Agriculture, Directorate of Economics and Statistics on crop production statistics and crop farm harvest prices. By matching district production levels to farm harvest prices by crop, I am able to construct an index of crop yields similar to that in Jayachandran (2006) for crop years 2002-10. The use of an index circumvents certain concerns arising from differences in crop suitability across districts.

Data on manufacturing enterprises are from the Annual Survey of Industries, reported annually for registered firms. Measures from enterprises with fewer than 100 employees are taken from a 20 per cent sample of firms representative at the state level. The ASI data used in this analysis do not report the district of the enterprise. As described in the empirical strategy section above, I adjust for the level of the data being broader than the level of the reform so as to best capture the spirit of the RD design.

District level data on several measures of interest, local GDP for example, are unavailable or available only sporadically. To overcome the lack of traditional measures, I consider data recorded from remote sensing on rainfall and the amount of light emitted at night from the TRMM satellite and DMSP-OLS Nighttime Lights Time Series, respectively. The night-time light data are used to proxy for changes in local GDP, as prescribed in Henderson et al. (2012). See the Data Appendix for greater detail on all data used in the analysis.

6 Results

In this section, I show that the reform resulted in a larger expansion of private sector bank branches in treated districts near the cutoff. I then confirm that this expansion was accompanied by expanding credit, detailing particular sectors of interest. I contrast the large discontinuity observed from the private sector banks to the null response from the public sector around the threshold, confirming predictions from the theoretical framework. I then turn to real measures of agricultural outcomes, showing responses in yields and output of important crops. I then show positive responses in manufacturing and conclude by showing faster growth in income using a proxy from remote sensing.

6.1 Banking

To motivate the primary set of empirical results, I present a before and after visual example from two years. Figure 5 presents a standard visual RD for operating private sector bank branches for the pre-reform year 2000 and the post-reform year 2012. The y-axis denotes the number of operating private bank branches per district, with dots reporting the local averages of districts falling within 200 persons per branch non-overlapping bins. The horizontal axis is the forcing variable of district population per branch centred on the national average and scaled to thousands of persons per branch. Considering the pre-reform year, districts do not appear to vary systematically in their number of branches. In the post-reform year, under-banked districts show higher numbers of operating branches relative to banked branches just on the other side of the cutoff. The discontinuity of the number of branches estimated at the cutoff from either side yields the local average treatment effect of the reform on private branches.

I now present results estimating equation 1 separately for each year, with operating private branches as the dependent variable.[24] I plot the intercept points at the cutoff from each annual local linear regression by year. The dashed line provides the estimated intercept from approaching the threshold along the under-banked side as in the classic RD graphical representation. The solid line reports the corresponding intercept approaching from the banked side. The vertical distance between the two, reported for each year, corresponds to the discontinuity at the cutoff estimated as τ in equation 1. A vertical line between the two points indicates a discontinuity with statistical significance at least at the 10 per cent level.[25]

The policy effect clearly emerges after 2006, shown in the right panel of Figure 6. Steadily higher branch growth in under-banked districts produces expanding positive discontinuities in

[24] Recall, districts maintain the same value of the forcing variable across years, so the set of districts remains unchanged. New districts since 2001 that claimed territory from more than one source district are dropped along with the source districts in all years. In addition, Thane and Pune districts in Maharashtra are dropped in all years, as is Varanasi district in Uttar Pradesh after 2002. See the Data Appendix for details.

[25] Thanks to Johannes Schmieder for helpful suggestions in clearly displaying the dynamic nature of the effect graphically. These figures not only present the average treatment effects, but place the level of the intercepts vertically so that the scale of the effect, as well as overall trends in growth, may be easily recognised. Note that these figures rely on estimation using a uniform kernel.

the average number of operating private sector bank branches. In contrast, the years leading up to the reform show little change in branching presence. The lack of pre-reform changes in the discontinuities provides a partial validation test of the randomisation of districts around the cutoff. The muted response in 2006 and 2007 is consistent with most banks making use of the policy grace period and waiting to submit their first ABEP until mid-2006. Those branches would then open towards the end of their licensing period in 2007. The estimates from the fuzzy RD with a triangular kernel presented in Table 4 verify that the largest discontinuities begin in 2008, estimated precisely at the 5 per cent and 1 per cent confidence levels.

The results for branch licences that have been granted, but may not yet represent an operating branch, support this interpretation of the branching pattern. The jump in the discontinuity magnitude, and the first year of statistical significance, occurs in 2007, shown in the top panel of Table 4. That is one year earlier than branches. Comparing branches with licences, the discontinuity from licences generally leads that on branches by one year for 2007-10. After 2010, the magnitudes of the discontinuities are generally in sync, consistent with the December 2009 reform removing the pre-approval requirement for branches opened in population centres with fewer than 50,000 people.

The response observed in operating branches and the corresponding timing of changes in licences, combined with the pre-reform null effects, provides strong evidence of a causal effect from the reform on expanding the presence of private sector banks in under-banked districts. The cumulative average effect of the policy in 2012 is estimated at approximately 10.6 more private sector branches in under-banked districts at the cutoff relative to the banked districts. The effect is a little more than 50 per cent of the sample mean reported in the table for 2012 at 20 private sector branches in districts around the cutoff. The size of the private sector presence increased for the sample overall in this time from an average of 10 branches per district in 2006 to 20 in 2012.

While the above analysis examines branching patterns, the effect on bank company presence can also be examined. Figure 7 and the corresponding Table 4 shows additional bank entry in the post-reform period of roughly one additional bank operating per district in the treated districts. The maximum estimated effect of nearly 1.5 additional banks is found in 2011. This effect is set against the average number of private sector banking companies for districts near the cutoff, which grew from 2.8 in 2002 to 3.5 in 2005 and ultimately 6.3 in 2012. These numbers likely underestimate the actual entry by new banking companies, as a series of mergers in the private banking sector occurred throughout the decade.

6.1.1 *Credit*

The mechanism through which the 2005 policy reform impacts lending behaviour is less direct. The reform cites opening branches in under-banked districts as a condition affecting total permissions to a bank. However, other conditions, including priority sector lending requirements, apply at the bank level rather than by district. Thus, the reform generates little direct pressure on bank credit and deposit activity. The theoretical framework in section 3, though, predicts that the anticipation of future competition would induce profit maximising banks to preemptively expand their credit in under-banked areas.

Figure 8 shows a pattern of expanding discontinuities in credit consistent with the hypothesis from the theoretical framework. The figure presents annual discontinuities in total district credit from private banks, analogous to the figure presented for operating branches discussed above. Prior to the policy announcement, the average level of credit around the threshold appeared similar. I broadly fail to reject the null hypothesis of continuity in the number of credit accounts in thousands and credit amounts outstanding in millions of rupees at the banked and under-banked districts' cutoff. In the post-reform period, positive discontinuities of significant magnitude are observed. The estimates for the number of accounts are estimated with precision at the 5 per cent level in 2007 and 2010, and the 10 per cent level for the other years in 2006-11, reported in Table 5.

The positive and statistically significant estimates of the discontinuity in credit accounts for 2006 and 2007 is consistent with banks responding preemptively to raised expectations of entry in under-banked districts following the policy implementation. The expansion is described as a preemptive response because it leads the positive discontinuities in bank branching first estimated with statistical precision in 2008. Taking the discontinuity in 2006 as the measure of this preemption, an estimated 6,220 additional credit accounts in the under-banked districts at the cutoff arose, which is 48 per cent of the sample mean for districts around the cutoff.[26] The greatest discontinuity was estimated in 2007, before significant branch entry, but following an estimated 2.4 additional increase in licences for branches in the underbanked area. The smaller positive discontinuities in 2005 come after the Vyas committee commissioning at the end of 2003 and the release of its report the following year, as discussed in the policy section. Accompanying these early changes in credit is a change in the composition of the banks in these districts, with fast growing banks expanding in districts as more inert banks were acquired by nationalised banks. This behaviour is consistent with aggressively growing banks acting preemptively on the expectation of reforms by expanding in areas likely to be more heavily contested in the future.

Lending activity to several sectors (for example, retail and whole sale trade, construction, mining and so on) compose the aggregate measures of credit. I narrow the focus here to credit for direct agriculture and personal loans. These sectors constitute a major portion of private sector bank business and are likely to exert a direct impact on households. The discontinuities estimated from this refined set, reported in the bottom two panels of Table 5, reflect the findings from the aggregate measures. Further, positive and sizeable discontinuities on credit amounts are now precisely measured at the 5 per cent and 1 per cent levels for years 2005-07 and 2010-11. The positive estimated discontinuity in 2004 that is significant at the 10 per cent level is not consistent with the framework. The magnitude is much smaller for this year, however, and may be a result of unrelated merger activity around that time.

[26] The list of under-banked districts was officially released in September 2005, making the 2006 credit measure the first following the realised list, while providing banks the least time to respond through branching. The estimated average treatment effects and sample means around the cutoff refer to the private sector banks.

6.1.2 *Public sector banks*

The result of a preemptive response in credit, derived in the theoretical framework, hinges on the profit maximising behaviour of a bank to secure a share of clients from which it may extract positive profits in the following period of fierce competition. Public sector banks, which follow less clear objective functions, are not predicted to show the same response as private sector banks. In fact, public sector banks do not show any response around the time of the reform. Further, no discontinuity of significant magnitude is estimated during the reform years. Table 6 shows the estimated effect of the reform on public sector credit accounts and amounts, which are imprecisely measured and small in magnitude relative to the sample mean for these outcomes in districts around the cutoff, shown in the last row of the table panels.

Public sector banks also exhibit a weak response in bank branching to the policy reform. In unreported results, discontinuities in public sector bank branches are small in magnitude relative to their existing stock and very imprecisely estimated. In behaviour that may reflect the influence of the NREGA programme, growth in public sector branching occurs on both sides of the cutoff at a similar rate beginning in 2007. The non-differential impact of NREGA around the cutoff will be discussed further in section 6.5. A positive discontinuity in branches appears in later years, though the small average treatment effects estimated for both branches and credit from the public sector suggests this type of policy reform may be ineffective for non-private sector entities.

6.2 Agriculture

Agriculture constitutes the primary economic activity for the majority of Indians. The 2001 population census reports that over 56 per cent of India's workers were engaged in agricultural or related activities at the time of the census which, due to the exclusion of marginal workers, likely provides a lower bound. Given that a concern over the availability of credit in rural areas motivated the commissioning of the Vyas Committee, which in turn catalysed a reform to rural branching and presumably the broader policy reform in 2005, I conduct here a closer examination of effects in agriculture. The amount of lending to agricultural purposes in less populated areas is shown to increase with the reform. Attention is then turned to the effect of expanded banking services on agricultural performance.

6.2.1 *Credit to agriculture*

Figure 10 shows the reduced form RD in the district percentage change in credit amount to rural and semi-urban areas from their 2001 levels, broken down by direct and indirect agricultural loans.[27] A positive and statistically significant response in under-banked districts is first detected for credit to direct agricultural activities in 2005, the year following the Vyas Committee Report. The effect grows in 2006 and diminishes slightly in 2007. A strong effect emerges in 2009

[27] The percentage change is approximated using the difference in logs of credit amounts from the 2001 reported levels.

and holds through the end of the sample period. The magnitude of higher growth in credit in treated districts is large. The average treatment effect exceeds the local means in 2005 and 2006, exceeding 60 per cent of the sample mean in all post-reform years. A strong effect from indirect agricultural loans emerges in 2009 as well, though with no evidence of a response to the reform prior to that time. All estimates are reported in Table 7.

The expansion of credit beginning in 2005 is consistent with the timing of the Vyas Committee and emphasis placed on agricultural lending by policy-makers, as well as the competition effects discussed above. The results from direct agricultural loans are interpreted as an initial increase due to the reform, followed by additional growth after 2008 potentially due to a variety of causes. The slowed growth after 2006 may be attributable to banks learning that the branching policy reform was less directly tied to agricultural lending than initially anticipated. Alternatively, a subsidised credit programme to farmers commencing around that time, exclusively administered through public sector banks, may have drawn away demand for private loans. The loss of demand may have washed out the private bank effect in direct agricultural credit.[28]

The growth in both forms of agricultural lending after 2008 in under-banked districts might be explained directly by the reform, or through an intersection of the reform and other policy refinements. New branches opening as a result of the policy reform ramped up during these years. A refinement to the branching policy in 2008 created greater incentive for banks to branch out into lower populated areas. The adoption of a new branching policy at the end of 2009 reduced the cost of branch entry in the rural and semi-urban areas of all districts, although it created additional incentives for under-banked districts. Each of these explanations focuses on the extensive expansion of branches.

Alternatively, a reform to priority sector lending in 2007 also placed greater emphasis on agricultural lending. Required investment in the Rural Development Infrastructure Fund for failing to meet priority sector quotas, first coming due in the 2009 financial year, was accounted as indirect agricultural lending by banks. A government-financed debt forgiveness scheme across all commercial banks in June 2008, for all delinquent direct-to-agriculture loans held by small farmers, may have reallocated new debt free borrowers across public and private sectors.[29] Lastly, the categorisation of loans by the RBI was revised in 2008, making direct comparisons by sector pre- and post-2008 less accurate. Without finer data on loans, disentangling the exact causes is not likely to be possible.

6.2.2 *Agricultural outcomes*

I find statistically significant results with economically meaningful magnitudes on individual crop yields and outputs that are consistent with a causal effect of credit expansion on agricultural

[28]The Credit Subvention Scheme operated through NABARD, and exclusively distributed through government sector banks, was initiated in 2006-07.

[29]Forgiven debtors became eligible for new loans, potentially resulting in some switching to private sector banks in those districts with greater branch coverage. This may also have contributed to the effect observed on indirect agricultural loans after 2008. Indirect-to-agriculture loans were excluded from the forgiveness scheme, potentially causing private sector banks to favour them in subsequent years.

outcomes. Considering crops individually, however, and absent price data for the crop output, complicates the interpretation of the results. Not every district produces each crop, or is well suited for every type of agriculture.[30] Therefore I relegate individual crop analysis to the online appendix and focus the discussion on an index of crop yields.

To construct a measure incorporating multiple crops and price data, I compute an index of crop yields similar to that used in Jayachandran (2006). The index is constructed as a weighted average of crop yields for rice, wheat, jowar, groundnut and cotton,[31] using individual crop revenue shares specific to the district as weights (see Data Appendix for details). I am able to construct the measure for the July-June years 2001-02 to 2009-10 from data on crop prices and production statistics collected at the district level. The price data for crops are available for a slightly smaller set of districts and generally restricted to crops for which the particular district produces greater volumes. The index carries the added benefit, however, that a wider set of districts in India produce at least one of the crops in volume, meaning the set of districts through time will change less than considering output from a single crop. The results from the reduced form RD analysis are shown in the top panel of Table 8. The estimates show positive discontinuities of sizeable magnitude beginning in 2005, though are estimated imprecisely except for 2009, and the bias-corrected estimates in 2005, 2009 and 2010.

To estimate the effect of banking activity on average crop yield, I estimate a fuzzy RD of the crop yield index on total private sector credit accounts, instrumenting for credit accounts with the discontinuity. In the bottom panel, I present the fuzzy RD results for the pre-reform and post-reform periods, pooling data across years and adding year fixed effects. No effect is estimated in the pre-reform period. In the post-reform period, I estimate an average effect of 0.023, with statistical significance at the 10 per cent level. The estimate may be interpreted as every thousand private bank accounts increases the crop yield by an average of 2.3 per cent. This is a little less than one third of the average effect of a positive rainfall shock, where rainfall is above the 80th percentile for that district, on crop yield estimated in Jayachandran (2006).[32]

[30]Many crops yield null results. Farmers may be moving in or out of crops based on anticipated prices. Yields of a popular crop may decrease if farmers expand into plots of land poorly conditioned for that crop. Alternatively, yields may increase if farmers invest more in existing crops when they are in high demand.

[31]The index in Jayachandran (2006) included sugarcane rather than cotton. The output and price data for sugarcane in my dataset contain many missing values, exhibit what appears to be rounding in several instances, and appear to report values for raw sugar rather than sugarcane at times, without always noting the distinction. For these reasons, and the strong observed effect on cotton, I substitute it for sugarcane in the index.

[32]The magnitude of the effect during the reform period varies depending on the choice of instrument. If accounts for direct agricultural and personal loans are used instead, then the effect is around 3 per cent, a little less than half of the effect found in Jayachandran (2006). Alternatively, leaving cotton out of the index reduces the effect in the post-reform period to about 1 per cent and loses statistical significance. In unreported results from replicating analysis with a differences-in-differences analysis limiting the sample of districts to those around the threshold, positive average effects of the policy on the crop index, and on total revenue from crops are found with statistical significance.

6.3 Industrial activities

Though the initial drive of the policy reform may have been to increase financial inclusion in low population areas and increase the credit flow to agriculture, many of the populated centres of under-banked districts benefited from increased branch entry. This section investigates to what extent manufacturing enterprises benefited from the expanded bank presence.

6.3.1 *Credit to manufacturing and processing*

Figure 11 presents the reduced form RD effect for the percentage change in credit amount to manufacturing and processing. The effect after 2007 resembles the expansion of bank branches, with a steadily growing positive effect in under-banked districts. The figure shows that the amount of credit exhibited little change in growth from the banked side at the cutoff after 2002. Unmatched growth from the under-banked side emerged beginning in 2008, with that difference doubling by 2010. The results from the fuzzy RD analysis presented in Table 9 confirm these findings, with statistically significant estimated effects in 2009 and 2010. The timing of the effect is consistent with the actual opening of branches. While it is possible that branches were lending to manufacturing at capacity at the time of the reform, such that no preemptive response was possible, it might also be the case that loans to manufacturing prior to 2009 may have been recorded by the banks under other categories, such as personal loans or other. The next section examines input decisions reported by registered manufacturing firms, including financing.

6.3.2 *Evidence from the ASI*

In Table 10 I present the results from difference-in-differences analysis using data from the ASI. The analysis uses years 1999-2010. In column (1) I estimate the effect on logged assets excluding land and inventory. The average treatment effect is positive but imprecisely estimated at a value of 17 per cent. The effect on logged working capital, in column (2), is estimated at 0.264 with significance at the 10 per cent level. The effect on the amount of outstanding loans held by the firm in column (3) is estimated to increase by 24 per cent with statistical significance at the 10 per cent level. Total investment presented in column (4) increased by 19.7 per cent, with statistical significance at the 10 per cent level. In the last column, the capital-labour ratio is estimated to increase by 3.4 in response to the policy and is also estimated with precision at the 10 per cent significance level. The sample mean of the capital-labour ratio for the under-banked states sample was 10.88 post-reform, making this a sizeable effect. The estimates are quite robust to considering other ranges of years around the reform. In each regression I control for the rural status of the enterprise, the age of the plant as measured by years since opening, the number of total enterprises in the firm to which the enterprise belongs, the logged number of employees at the enterprise to control for size, and state fixed effects with state-specific time trends. I exclude industry fixed effects as new NIC codes were adopted in 2008, potentially making some industry codings inconsistent through the time series. In practice, the inclusion of three-digit NIC codes has little effect on the estimates.

The estimates are consistent with the expansion of the banking sector having a significant impact on manufacturing. The significant increase in loans carried by enterprises from under-banked districts in the post-reform years would indicate that the increased banking activity is finding its way to the industrial sector. The increases in working capital as well as total investments suggests firms are expanding the use of productive inputs with the expansion of credit. Further, the increase in the capital-labour ratio is consistent with previously credit constrained firms making investments in capital as those constraints are relaxed with the inflow of new formal credit. These adjustments to the productive technologies of the firm are likely to result in changes in efficiency. If credit rationing resulted in the misallocation of credit, the expansion of credit may produce large impacts if it helps correct inefficient dispersions of marginal products of capital across firms.

6.4 Economic growth and light emitted at night

I return to the RD design for the final analysis examining the effect of banking expansion on overall economic growth at the district level. Henderson et al. (2012) established that changes to the amount of light emitted at night provide a reliable proxy for economic growth under certain caveats.[33] This analysis accounts for these concerns by estimating the discontinuity in the difference of logged average district light since 2004. Thus, the dependent variable can be interpreted as the approximate percentage change in average light emitted in a location from its 2004 baseline level. The RD compares these changes in estimating the discontinuity at the threshold.

Figure 12 graphically reports the discontinuities estimated using a second degree polynomial, which better captures the underlying data. Since the level of light is reported from measurements taken during the calendar year, 2005 is the first year with months under the enacted reform. Estimates are presented in Table 11. A slight negative discontinuity is estimated in the first year and is a small fraction of the average percentage change in light for districts in the sample. The discontinuity is small again but positive in 2006, though the average change in districts was negative overall. A positive jump in the discontinuity to 11.3 per cent appears in 2007 and is estimated significantly at the 1 per cent level, with the average change in light for districts in the sample increasing as well to 11.4 per cent. A similar response is found in 2008 though the average growth in light from 2004 declined, such that the relative magnitude is greater. Lower levels of light are emitted overall in 2009. The last three years show similar discontinuities in light to 2007 and 2008, with 2011 estimated with precision at the 10 per cent confidence level.

To estimate the effect of expanding branch presence on overall economic growth in districts, I perform a fuzzy RD of the change in light on private bank branches for the pre-reform period, which in this case is only 2005, and the post-reform period constituting years 2006-12 pooled together controlling for year fixed effects in addition to district population and its square. I run the estimation using local linear regressions because these better fit the bank branching data

[33]Important among these is the prescription to compare changes in light through time for one area to those in another, rather than comparing levels of light only across places or levels of light only across time. Such comparisons can help account for the switching of satellites, ageing of instruments, and differences in the processing of very low levels of light across years.

and offer a strong first stage. The pre-reform effect reported in the lower panel is negative and small, consistent with the reduced form estimate for 2005. The conventional estimate reported in the post-reform column is estimated as positive, but small and insignificant. However, the bias-corrected measure which accounts for a local quadratic estimate with a wider bandwidth, better capturing the quadratic relationship in the night light data, yields a positive and significant coefficient. This estimate is significant at the 1 per cent level and has a value of 0.0115. The coefficient may be interpreted as the average effect of a bank branch during the reform period was to increase night-time light by 1.15 per cent. Taking the estimated elasticity of night-time light to GDP from Henderson et al. (2012) of 0.3, this implies that each bank branch raises local GDP by approximately 0.33 per cent. The average increase in bank branches in the post-reform period is estimated at approximately 5, implying the total effect was an average increase of local GDP in the districts by 1.65 per cent.

6.5 Robustness and discussion

6.5.1 *Robustness to NREGA*

A competing explanation for the change in the spatial allocation of bank branches, increased banking activity, and subsequent responses in economic outcomes is the introduction of the Mahatma Gandhi National Rural Employment Guarantee Act (NREGA) that coincided in time with the branching policy reform. The act constitutes a public works programme aimed at relieving poverty in rural areas by providing 100 days of guaranteed work to individuals from rural areas. The implementation of NREGA occurred in three stages, with 200 districts selected to begin the programme in the fiscal year from April 2006 to March 2007, with 130 new districts introduced in 2007-08 and the remaining 263 districts introduced in 2008-09. Zimmermann (2012) and Klonner and Oldiges (2014) analyse the effect of NREGA using these rollout phases and provide background on the programme. Of particular importance to the current analysis, NREGA benefits were distributed through bank accounts.[34] One may conclude that this would increase the demand for formal banking, potentially increasing both the geographic reach and level of banking services. While likely true, to confound the current results there must also be a discontinuous break in the implementation of the programme and disbursement of benefits at the "under-banked" cutoff used for the regression discontinuity.

Districts were assigned to the various roll-out phases based on a composite index on district "backwardness" from the National Planning Commission (2003). In Table 12 I test whether a discontinuity in phase assignment can be detected at the cutoff. A significant discontinuity would suggest a correlation with the NREGA programme. The test fails to reject the null hypothesis of continuity at the cutoff for all three phases. Thus, the NREGA phase assignment, and therefore likely its benefits as well, would be unexpected to differ at the cutoff.[35]

[34]NREGA benefits were primarily disbursed through public sector banks and post office bank accounts. Private sector banks did not receive general authorisation to disburse NREGA funds until 31st January 2012 (Ministry of Rural Development, 2012).

[35]In analysis not shown, I perform a visual RD of the district composite index at the under-banked cutoff. No discontinuity is observable at the cutoff. Further, the general notion that persons per branch is generally increasing

6.5.2 *Behaviour at the cutoff*

The theoretical framework suggests banks may face an incentive to reduce investments in untreated districts near the cutoff. The post-reform branching stock in Figure 5 may show a steeper negative slope just to the left of the cutoff than further into the set of control districts. While that pattern may indicate a particularly large response in the first stage, it does not constitute a threat in itself to the estimates in the fuzzy RD analysis. The variation in banking assets, branches or accounts remains driven by the reform under the assumptions of the estimation strategy. Further, the effect on banking outcomes is observable and included as part of the fuzzy RD.[36]

with worsening district conditions is confirmed by the trend of the index on "backwardness". Out of concern that the omitted districts are disproportionately from one side of the cutoff or the other, I repeat the McCrary test only including districts missing the composite index value. I fail to reject the null hypothesis of continuity in the density of districts at the cutoff with the discontinuity estimate in the log difference in height at -31 and a standard error of 38.

[36]A concern for identifying the unbiased relationship between banking and economic outcomes could come from potential general equilibrium effects of the reform. For instance, improved agricultural output in treated districts from better access to finance may affect markets in geographically nearby districts otherwise not impacted by the reform. Such impacts that feed back into local credit markets may lead to biased estimates. The potential redistribution of assets following the reform precludes a full discussion of welfare effects. Estimating the general equilibrium effects through a fully specified model is left to future work.

7 Conclusions

This paper analyses a previously unstudied policy reform in India introduced in 2005, finding new evidence of positive effects from bank branch expansion on economic growth. The concentrated response to the reform from private sector banks highlights that banks and their branches act as strategic players responding to incentives. The mobilisation of the private banking sector helps explain the positive findings in this work, and bridge the gap between the literature promoting financial inclusion and the frequently null results from micro empirical analyses examining other financial interventions in developing country settings.

Importantly, credit expansion and its effects do not appear to have been confined to urban areas, a common concern in developing countries. Rural and semi-urban markets in under-served areas also exhibited increases of credit from private sector banks. Agricultural productivity and the capital intensity of manufacturing are shown to increase in areas receiving higher credit due to the reform. I estimate that an increase of 1,000 private bank credit accounts in a district raises average crop yield by 2.3 per cent. This effect is a little less than one third of the effect Jayachandran (2006) measures on crop yield from positive rainfall shocks. Manufacturing enterprises in areas with increased access to banking exhibited higher growth in total investments, working capital and capital labour ratios. The empirical strategy in my paper identifies these effects independently of growth from the NREGA public work programme introduced around this time, suggesting the expansion of credit as a complementary source of agricultural and industrial growth. I confirm the aggregate effect on local GDP growth using night-time light intensity data, estimating that each additional private bank branch led to a 0.33 per cent increase in local GDP.

The results have implications for broader areas. With respect to growth, the role of banking in facilitating the link between improved agricultural productivity and industrialisation, as examined in recent work by Bustos et al. (2016); Santangelo (2016); Asher and Novosad (2012) requires further study. Beyond redistributing productivity gains across sectors, the findings in this paper provide evidence that banking access can generate direct growth in productive sectors. Second, further research into the efficient expansion of bank branches and bank access is required. Policies aiming to direct branch openings in specific areas can distort the distribution of resources. Bank access within communities, to both deposits and credit, may be uneven across land owners and labourers. Recent work by Mobarak and Rosenzweig (2014) shows that uneven access to instruments helping to mitigate risk can result in adverse welfare outcomes in some instances. Future work should address the issues of aggregate efficiency and inequality following expanded bank access.

References

(2005). Master circular on branch authorisation. Technical report, Reserve Bank of India.

Allen, T. and Atkin, D. (2015). Volatility, insurance, and the gains from trade. Technical report, mimeo.

Angrist, J. D. and Lavy, V. (1999). Using maimonides' rule to estimate the effect of class size on scholastic achievement. *The Quarterly Journal of Economics*, 114(2):533–575.

Asher, S. and Novosad, P. (2012). Factory farming: The impact of agricultural output on local economic activity in india. Technical report, Working Paper.

Banerjee, A. and Duflo, E. (2001). The nature of credit constraints: Evidence from an indian bank. In *Manuscript, MIT. Available at: http://www.chicagobooth.edu/research/workshops/AppliedEcon/archive/WebArchive20012002/duflo.pdf*.

Banerjee, A., Duflo, E., Glennerster, R., and Kinnan, C. (2015). The miracle of microfinance? evidence from a randomized evaluation. *American Economic Journal: Applied Economics*, 7(1):22–53.

Banerjee, A. V., Cole, S., and Duflo, E. (2004). Banking reform in india. *India Policy Forum*, 1(1):277–332.

Banerjee, A. V. and Duflo, E. (2014). Do firms want to borrow more? testing credit constraints using a directed lending program. *The Review of Economic Studies*, 81(2):572–607.

Barone, G., Felici, R., and Pagnini, M. (2011). Switching costs in local credit markets. *International Journal of Industrial Organization*, 29(6):694–704.

Black, S. E. (1999). Do better schools matter? parental valuation of elementary education. *Quarterly journal of economics*, pages 577–599.

Burgess, R. and Pande, R. (2005). Do rural banks matter? evidence from the indian social banking experiment. *American Economic Review*, 95(3):780–795.

Bustos, P., Garber, G., and Ponticelli, J. (2016). Capital allocation across sectors: Evidence from a boom in agriculture. Technical report, Central Bank of Brazil, Research Department.

Calonico, S., Cattaneo, M. D., and Titiunik, R. (2014). Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica*, 82(6):2295–2326.

Cole, S. (2009). Financial development, bank ownership, and growth: Or, does quantity imply quality? *The Review of Economics and Statistics*, 91(1):33–51.

Conning, J. and Udry, C. (2007). Rural financial markets in developing countries. *Handbook of Agricultural Economics*, 3:2857–2908.

Dehejia, R. and Lleras-Muney, A. (2007). Financial development and pathways of growth: State branching and deposit insurance laws in the united states, 1900–1940. *Journal of Law and Economics*, 50(2):239–272.

Farrell, J. and Klemperer, P. (2007). Chapter 31 coordination and lock-in: Competition with switching costs and network effects. volume 3 of *Handbook of Industrial Organization*, pages 1967 – 2072. Elsevier.

Fulford, S. L. (2013). The effects of financial development in the short and long run: Theory and evidence from india. *Journal of Development Economics*, 104:56–72.

Gehrig, T. and Stenbacka, R. (2007). Information sharing and lending market competition with switching costs and poaching. *European Economic Review*, 51(1):77–99.

Gelman, A. and Imbens, G. (2014). Why high-order polynomials should not be used in regression discontinuity designs. Technical report, National Bureau of Economic Research.

Hahn, J., Todd, P., and Van der Klaauw, W. (2001). Identification and estimation of treatment effects with a regression-discontinuity design. *Econometrica*, 69(1):201–209.

Henderson, J. V., Storeygard, A., and Weil, D. N. (2012). Measuring economic growth from outer space. *American Economic Review*, 102(2):994–1028.

Hsieh, C.-T. and Klenow, P. J. (2009). Misallocation and manufacturing tfp in china and india. *The Quarterly Journal of Economics*, 124(4):1403–1448.

Imbens, G. and Kalyanaraman, K. (2011). Optimal bandwidth choice for the regression discontinuity estimator. *The Review of Economic Studies*, page rdr043.

Imbens, G. W. and Lemieux, T. (2008). Regression discontinuity designs: A guide to practice. *Journal of Econometrics*, 142(2):615–635.

Jayachandran, S. (2006). Selling labor low: Wage responses to productivity shocks in developing countries. *Journal of Political Economy*, 114(3):538–575.

Jayaratne, J. and Strahan, P. E. (1996). The finance-growth nexus: Evidence from bank branch deregulation. *The Quarterly Journal of Economics*, pages 639–670.

Kaboski, J. P. and Townsend, R. M. (2011). A structural evaluation of a large-scale quasi-experimental microfinance initiative. *Econometrica*, 79(5):1357–1406.

Kaboski, J. P. and Townsend, R. M. (2012). The impact of credit on village economies. *American Economic Journal. Applied Economics*, 4(2):98.

King, R. G. and Levine, R. (1993). Finance and growth: Schumpeter might be right. *The Quarterly Journal of Economics*, pages 717–737.

Klemperer, P. (1987). Markets with consumer switching costs. *The Quarterly Journal of Economics*, pages 375–394.

Klonner, S. and Oldiges, C. (2014). Safety net for india's poor or waste of public funds? poverty and welfare in the wake of the world's largest job guarantee program. Technical report, University of Heidelberg, Department of Economics.

Kochar, A. (2011). The distributive consequences of social banking: a microempirical analysis of the indian experience. *Economic Development and Cultural Change*, 59(2):251–280.

Krishnan, K., Nandy, D., and Puri, M. (2014). Does financing spur small business productivity? evidence from a natural experiment. *Review of Financial Studies*, page hhu087.

La Porta, R., Lopez-de Silanes, F., and Shleifer, A. (2002). Government ownership of banks. *The Journal of Finance*, 57(1):265–301.

Lee, D. S. (2008). Randomized experiments from non-random selection in us house elections. *Journal of Econometrics*, 142(2):675–697.

Lee, D. S. and Lemieux, T. (2010). Regression discontinuity designs in economics. *Journal of Economic Literature*, 48(2):281–355.

Lee, D. S., Moretti, E., and Butler, M. J. (2004). Do voters affect or elect policies? evidence from the u. s. house. *The Quarterly Journal of Economics*, 119(3):807–859.

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

Mobarak, A. M. and Rosenzweig, M. (2014). Risk, insurance and wages in general equilibrium. Technical report, National Bureau of Economic Research.

Panagariya, A. (2006). Bank branch expansion and poverty reduction: A comment. Technical report, Mimeo, Columbia University.

Rajan, R. G. and Zingales, L. (1998). Financial dependence and growth. *The American Economic Review*, 88(3):559–586.

RBI (2009). Report of the group to review branch authorisation policy.

Santangelo, G. (2016). Firms and farms: The impact of agricultural productivity on the local indian economy. Technical report, Working Paper.

Stiglitz, J. E. and Weiss, A. (1981). Credit rationing in markets with imperfect information. *The American Economic Review*, pages 393–410.

Thistlethwaite, D. L. and Campbell, D. T. (1960). Regression-discontinuity analysis: An alternative to the ex post facto experiment. *Journal of Educational Psychology*, 51(6):309.

Van der Klaauw, W. (2002). Estimating the effect of financial aid offers on college enrollment: A regression–discontinuity approach. *International Economic Review*, 43(4):1249–1287.

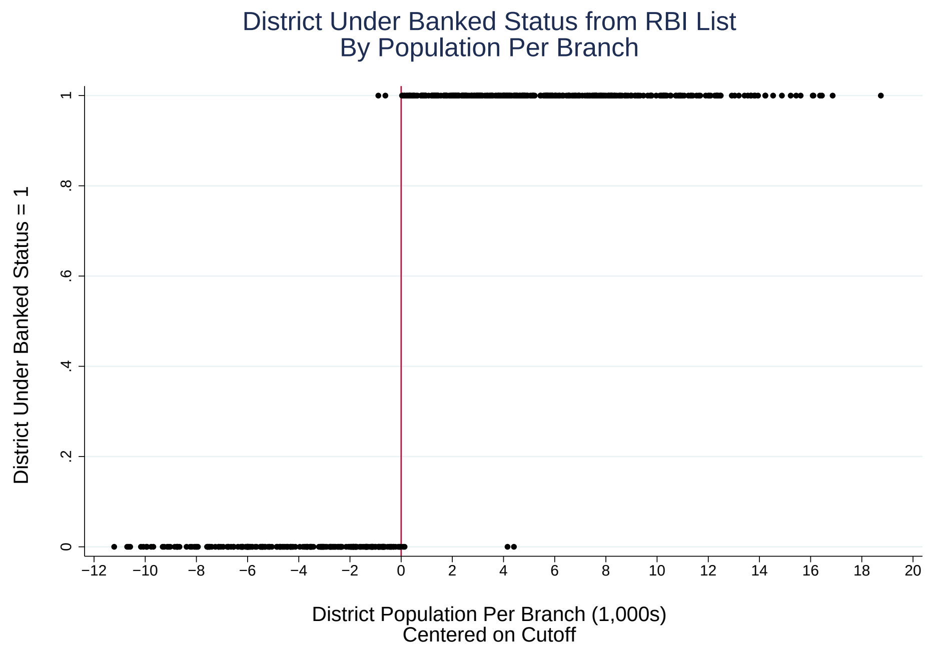
Vesala, T. (2007). Switching costs and relationship profits in bank lending. *Journal of Banking & Finance*, 31(2):477–493.

Zimmermann, L. (2012). Labor market impacts of a large-scale public works program: evidence from the indian employment guarantee scheme.

Appendix

Figures

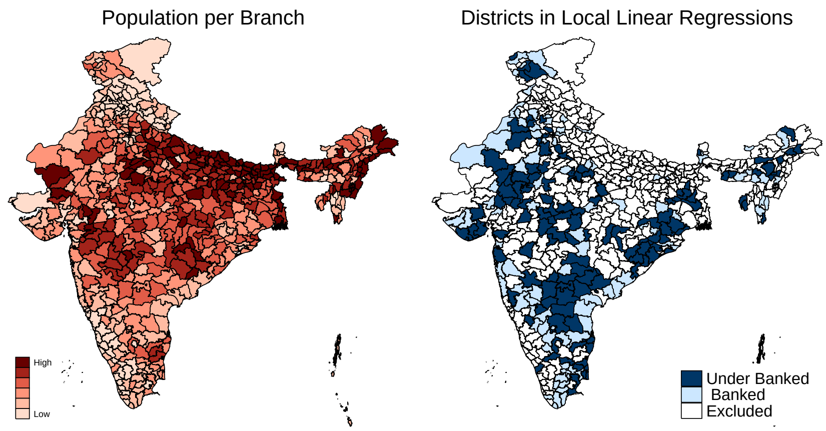
Figure 1: Under-banked status by district population per branch



Source: Population Census 2001, RBI and author's calculations.
Note: The dots report the under-banked status of a district, taking a value equal to 1 if the district appeared on the list of under-banked districts in the 2006 RBI MC on Branching Authorisation Policy, and zero otherwise. The forcing variable, district population per branch centred on the national average, is on the x-axis scaled to thousands of persons per branch. Values to the right of the cutoff are predicted to have under-banked status. 369 districts of 572 have under-banked status, with 5 incorrect predictions based on the rule.

Figure 2: Geographic distribution of district level population per branch

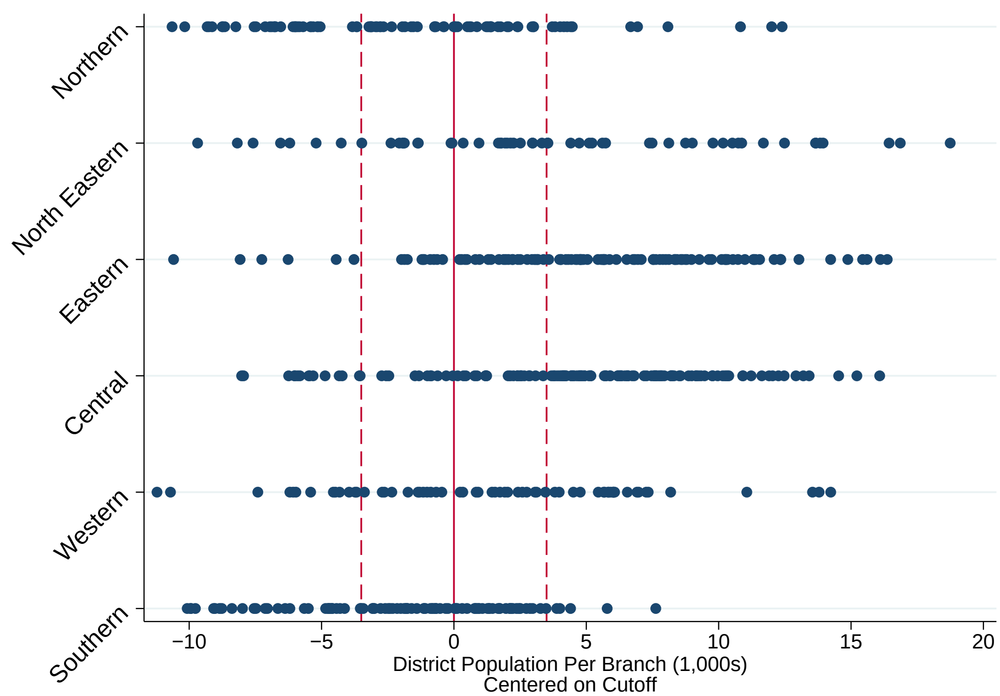
A. Maps of under-served areas by formal banking



Source: Population Census 2001, RBI and author's calculations.

Note: Heat map of district population per branch is on the left. District under-banked status, excluding districts outside local linear regressions bandwidth, is on the right.

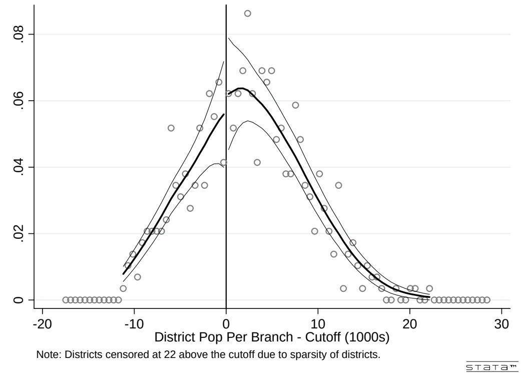
B. District population per branch across RBI regions



Source: Population Census 2001, RBI and author's calculations.

Note: Districts are plotted with a dot to indicate their population per branch, scaled to thousands with the threshold normalised to zero, along the horizontal axis while the vertical axis separates districts according to their region assigned by the RBI. A solid vertical line is drawn at the threshold, with dashed vertical lines indicating a bandwidth of 3,500 persons per branch, the same used throughout the analysis.

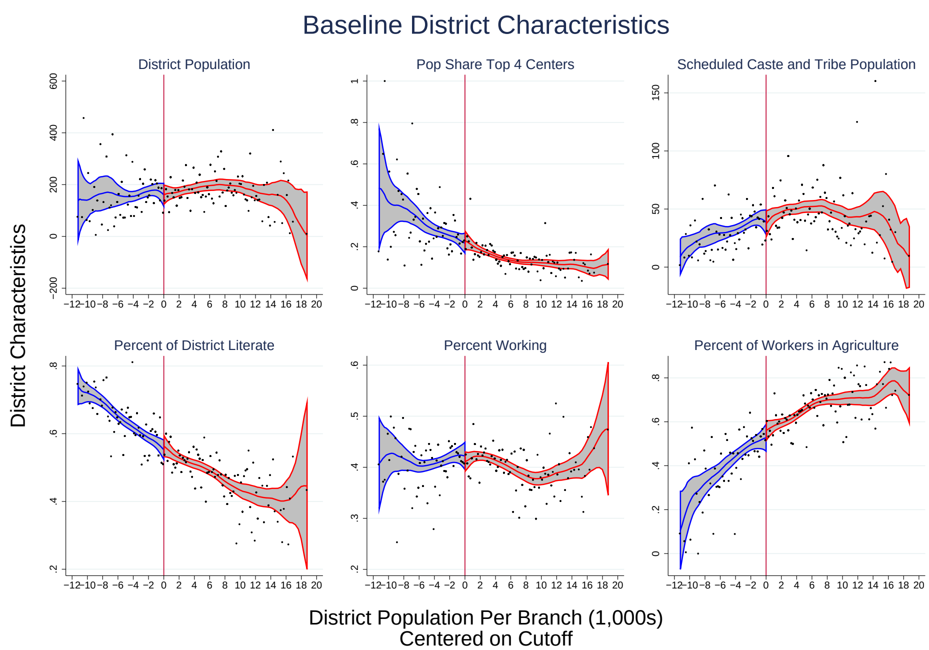
Figure 3: Visual McCrary test



Source: Population Census 2001, RBI and author's calculations.

Note: The graph plots a density of districts along the forcing variable, district population per branch, centred on the cutoff. The discontinuity estimate in the log difference in height is 6.6 with a standard error of 22. I fail to reject the null hypothesis of continuity at the cutoff, suggesting a lack of manipulation.

Figure 4: Continuity around the threshold



Source: Population Census 2001, RBI and author's calculations.

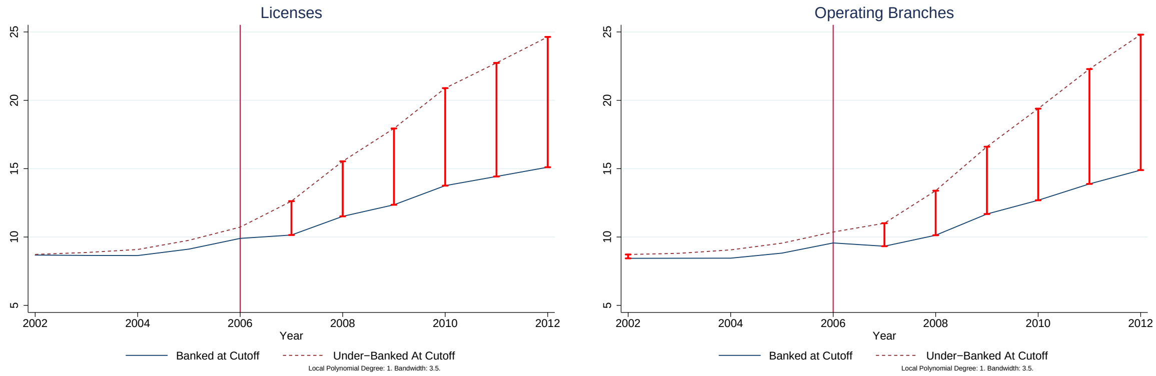
Note: The figure presents baseline district characteristics taken from the 2001 population census of India, with dots reporting local averages for districts falling within non-overlapping 200 persons per branch bins. The horizontal axis is the forcing variable of district population per branch centred on the cutoff. Districts predicted to have under-banked status fall to the right of the cutoff. The estimated y-value from a local linear regression of bandwidth 3,500 persons per branch is shown at each x-value, allowing for different slopes on either side of the cutoff, with 5 per cent confidence intervals.

Figure 5: Visual RD: operating private bank branches

Source: Population Census 2001, RBI and author's calculations.

Note: Each plot presents the number of operating private sector bank branches within a district, in respective years, with dots reporting local averages of branches for districts falling within non-overlapping 200 persons per branch bins. The horizontal axis is the forcing variable of district population per branch centered on the cutoff and scaled to thousands of persons per district. The estimated local linear regressions, with a 3,500 persons per district bandwidth and triangular kernel, at each x-value and the 5 per cent confidence intervals are shown, allowing for different slopes on either side of the cutoff. The year 2000 in the left plot shows a pre-reform example of branches around the cutoff. The figure on the right shows the cumulative effect of the policy on operating branches since its implementation in 2005. Local averages greater than 40 are not shown in the plots, but were included in local linear regressions.

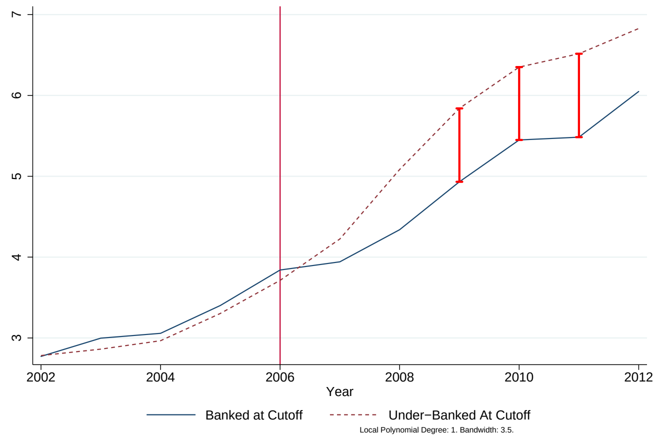
Figure 6: Discontinuity from reduced form: operating private bank branches



Source: Population Census 2001, RBI and author's calculations.

Note: Estimated using local linear regressions with controls for district population and its square, the per cent of workers in agriculture and the pre-randomisation 2001 value of the dependent variable. Bandwidths are set 3,500 persons per branch and estimated using a uniform kernel. The figure plots the estimated intercepts at the cutoff from the estimation of the RD equation repeated annually. The red dashed line provides the estimated intercept from approaching the threshold along the under-banked side. The solid blue line reports the corresponding intercept approaching from the banked side. The distance between the two, reported for each year, shows the estimated discontinuity at the threshold. A solid line between the two points indicates an estimated discontinuity with statistical significance of at least the 10 per cent level. The thin vertical line at 2006 represents the first estimation made following the reform.

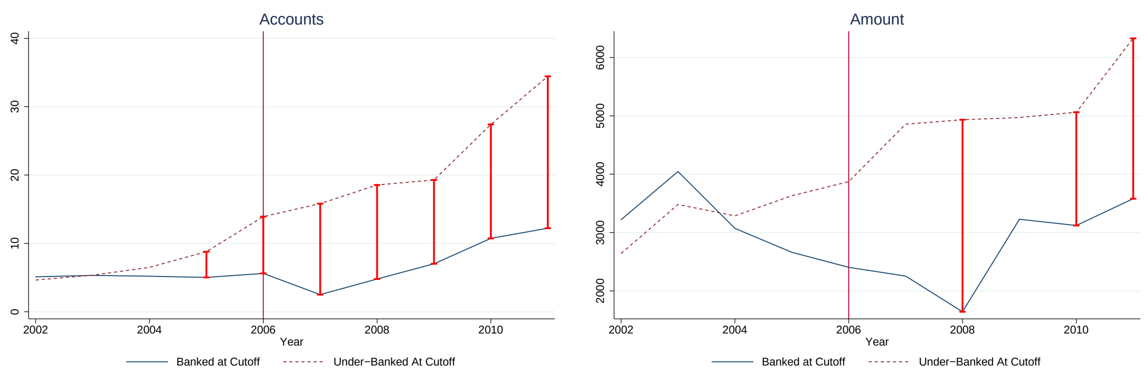
Figure 7: Discontinuity from reduced form: operating private banks



Source: Population Census 2001, RBI and author's calculations.

Note: Estimated using local linear regressions with controls for district population and its square, the per cent of workers in agriculture and the pre-randomisation 2001 value of the dependent variable. Bandwidths are set 3,500 persons per branch and estimated using a uniform kernel. See notes from Figure 6 for a detailed graph description.

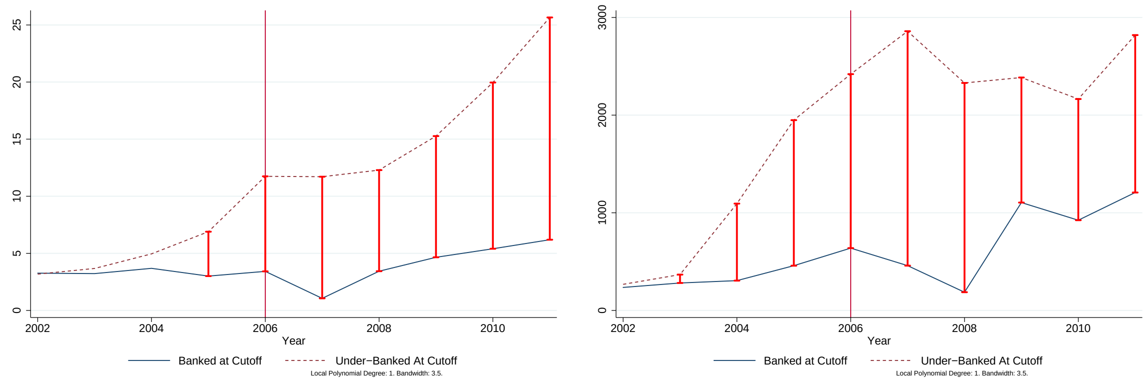
Figure 8: Discontinuity from reduced form: private banks aggregate credit



Source: Population Census 2001, RBI, TRMM and author's calculations.

Note: Accounts reported in thousands. Amounts reported in millions of rupees. Estimated using local linear regressions with controls for district population and its square, the per cent of workers in agriculture, a control for monsoon rainfall and the pre-randomisation 2001 value of the dependent variable. Bandwidths are set 3,500 persons per branch and estimated using a uniform kernel. See notes from Figure 6 for graph description.

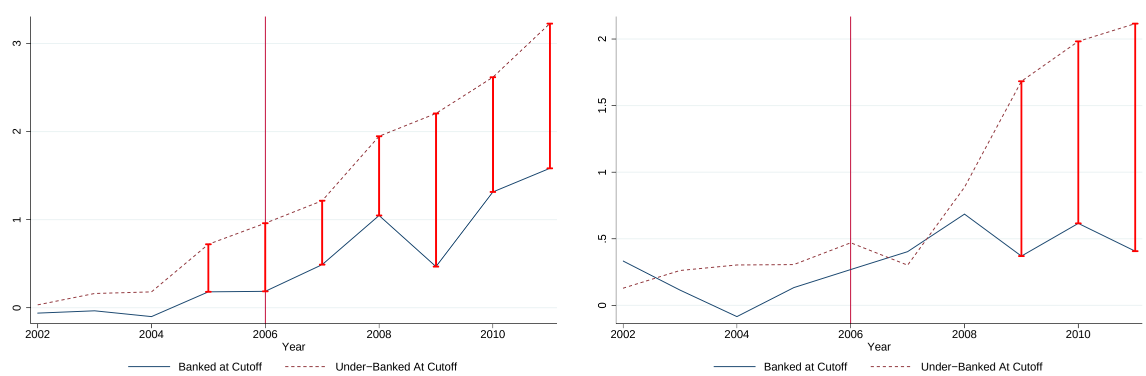
Figure 9: Discontinuity from reduced form: private credit to direct agriculture and personal loans



Source: Population Census 2001, RBI, TRMM and author's calculations.

Note: Accounts reported in thousands. Amounts reported in millions of rupees. Estimated using local linear regressions with controls for district population and its square, the per cent of workers in agriculture, a control for monsoon rainfall and the pre-randomisation 2001 value of the dependent variable. Bandwidths are set 3,500 persons per branch and estimated using a uniform kernel. See notes from Figure 6 for graph description.

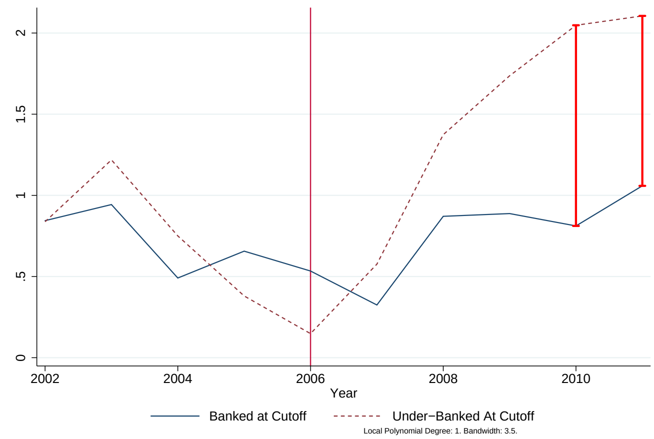
Figure 10: Discontinuity from reduced form: percentage change in private credit amount to agriculture in rural and semi-urban areas



Source: Population Census 2001, RBI, TRMM and author's calculations.

Note: Direct agriculture is on the left, indirect agriculture is on the right. Percentage change is approximated using difference in logs relative the value reported in 2001. Estimated using local linear regressions with controls for district population and its square, the per cent of workers in agriculture and a control for monsoon rainfall . Bandwidths are set 3,500 persons per branch and estimated using a uniform kernel. See notes from Figure 6 for graph description.

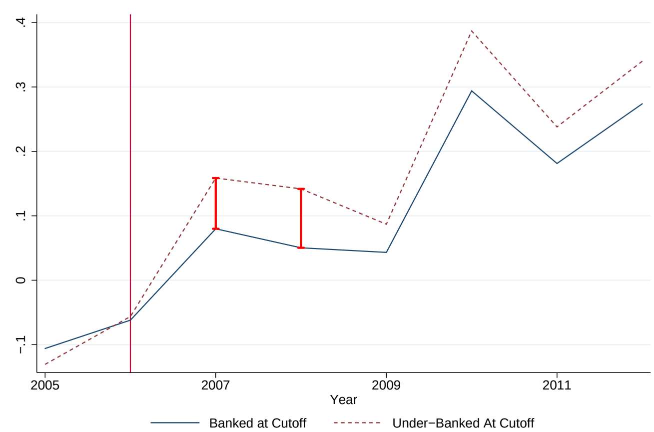
Figure 11: Discontinuity from reduced form: percentage change in private credit amount to manufacturing and processing from 2001 level



Source: Population Census 2001, RBI, TRMM and author's calculations.

Note: Percentage change is approximated using difference in logs relative the value reported in 2001. Estimated using local linear regressions with controls for district population and its square, the per cent of workers in agriculture and a control for monsoon rainfall. Bandwidths are set 3,500 persons per branch and estimated using a uniform kernel. See notes from Figure 6 for graph description.

Figure 12: Discontinuity from reduced form: difference in log mean district light from 2004 level



Source: Population Census 2001, NOAA, RBI, TRMM and author's calculations.

Note: Estimated using local quadratic regressions with controls for district population and its square, the per cent of workers in agriculture and a control for monsoon rainfall. Bandwidths are set to 3,500 persons per branch and estimated using a uniform kernel. See notes from Figure 6 for graph description.

Tables

Table 1: Continuity tests for baseline values at the cutoff

VARIABLES	(1) Population	(2) Pop Share Top 4 Centers	(3) Sched Caste Tribe Pop	(4) Pct Literate	(5) Pct Pop Working	(6) Share Workers in Agri	(7) PrivBranches2000
Conventional	0.839	0.0135	-1.436	0.0114	-0.0114	0.0321	0.192
	[35.38]	[0.0400]	[8.483]	[0.0219]	[0.0197]	[0.0482]	[3.026]
Bias-corrected	16.01	0.0218	0.265	0.0187	-0.0129	0.0363	0.567
	[35.38]	[0.0400]	[8.483]	[0.0219]	[0.0197]	[0.0482]	[3.026]
Robust	16.01	0.0218	0.265	0.0187	-0.0129	0.0363	0.567
	[42.75]	[0.0461]	[9.840]	[0.0261]	[0.0235]	[0.0591]	[3.527]
Bandwidth	3.500	3.500	3.500	3.500	3.500	3.500	3.500
N_Banked	95	95	95	95	95	95	95
N_UBanked	122	122	122	122	122	122	122
DepMean	176.7	0.221	45.24	0.553	0.421	0.550	7.198

Source: Population Census 2001, RBI and author's calculations.

Note: Estimated using local linear regressions with no controls. Bandwidths are set 3,500 persons per branch and estimated using a triangular kernel.

Table 2: Summary statistics

Banking

	Banked, Pre-reform			Banked, Post-reform			Under Banked, Pre-reform			Under Banked, Post-Reform		
	N	Mean	Std. Dev.	N	Mean	Std. Dev.	N	Mean	Std. Dev.	N	Mean	Std. Dev.
Branches												
SBI	610	28.618	23.095	732	33.238	27.971	900	21.35	16.381	1080	24.595	19.693
Nationalised	610	69.805	62.759	732	80.634	73.984	900	45.444	44.86	1080	51.54	50.432
RRB	610	21.523	21.684	732	23.001	22.171	900	28.221	22.147	1080	29.207	22.946
Foreign	610	0.121	0.624	732	0.243	1	900	0.018	0.199	1080	0.112	0.457
Old Private	610	11.807	16.582	732	11.628	15.295	900	4.198	9.298	1080	4.589	10.001
New Private	610	2.428	4.6	732	7.25	10.687	900	0.794	2.372	1080	4.049	6.154
Public Banks	610	120.154	87.587	732	137.926	105.531	900	95.064	66.491	1080	106.006	76.429
Private Banks	610	14.234	18.58	732	18.878	20.755	900	4.992	10.375	1080	8.638	13.96
Credit Amount												
SBI	610	5293.635	5980.068	732	11037.746	12248.838	900	3285.651	5986.45	1080	6507.465	8548.992
Nationalised	610	10236.988	13154.392	732	22228.233	33180.444	900	4602.575	5692.052	1080	9362.257	12494.337
RRB	610	870.748	1198.64	732	1738.277	2270.793	900	950.135	1134.78	1080	1869.281	2256.909
Foreign	610	201.344	727.787	732	487.36	1620.559	900	50.173	293.389	1080	191.788	1414.19
Private	610	3813.913	7071.325	732	7637.427	12055.826	900	1354.922	3542.466	1080	2437.963	5464.27
Credit Accounts												
SBI	610	30945.372	31517.419	732	47639.104	50181.875	900	24107.006	24218.304	1080	38046.444	39105.763
Nationalised	610	60582.561	60584.955	732	89278.02	97041.327	900	37963.999	38526.215	1080	55938.739	58976.202
RRB	610	22255.538	33920.327	732	30088.209	47295.116	900	28251.067	34646.607	1080	36354.233	48093.88
Foreign	610	134.425	772.631	732	319.858	1656.413	900	51.02	564.603	1080	119.098	874.722
Private	610	9792.657	14751.414	732	25507.242	35027.737	900	3214.418	7356.894	1080	9889.303	22363.595
Deposit Amount												
SBI	607	9599.797	10660.293	732	16412.707	20661.421	892	6104.533	6197.594	1078	10180.87	10886.087
Nationalised	607	20027.738	26126.927	732	33469.464	51159.493	892	9745.183	12975.665	1078	15306.32	20677.413
RRB	607	1340.932	1519.9	732	2212.508	2520.006	892	1807.669	1792.853	1078	2828.679	2818.4
Foreign	607	181.203	1207.168	732	611.752	4849.064	892	20.185	243.413	1078	65.089	603.547
Private	607	4695.24	8722.103	732	8973.14	17799.643	892	1371.376	2938.947	1078	2798.099	5257.67
Deposit Accounts												
SBI	607	203.438	178.676	732	298.246	276.023	892	147.726	130.511	1078	232.131	228.963
Nationalised	607	502.83	502.301	732	683.751	696.657	892	294.637	342.959	1078	410.485	464.146
RRB	607	76.55	101.221	732	118.76	157.796	892	100.515	109.819	1078	157.422	174.789
Foreign	607	0.98	6.606	732	2.268	14.987	892	0.188	2.342	1078	0.396	2.834
Private	607	91.003	124.465	732	136.977	184.145	892	30.155	63.394	1078	50.568	93.778

Source: RBI Master Office File, BSR 1 and BSR 2 years 2001-2011. Sample includes years 2001-2011 for districts falling within 5 thousand persons per branch of the national average. Each year includes 122 banked districts and 180 under banked districts, from a total of 572 districts considered. Amounts are reported in Rupees million adjusted to 2011q4 prices; Accounts are reported in thousands.

Table 3: Summary statistics continued...

Agriculture

	Banked, Pre-reform			Banked, Post-reform			Under Banked, Pre-reform			Under Banked, Post-reform		
	N	Mean	Std. Dev.	N	Mean	Std. Dev.	N	Mean	Std. Dev.	N	Mean	Std. Dev.
Cotton												
Area	403	32,656	53,321	349	31,406	56,677	619	31,351	64,876	471	37,076	75,472
Output	403	59,959	127,462	349	100,347	229,598	619	41,581	89,199	471	86,119	203,562
Productivity	403	1.61	0.98	349	2.12	1.38	619	1.35	0.84	471	1.55	1.28
Maize												
Area	560	11,945	20,923	470	15,124	28,518	968	16,400	32,962	761	16,688	36,426
Output	560	27,988	57,175	470	48,069	103,819	968	28,449	64,162	761	34,070	87,053
Productivity	560	1.87	1.19	470	2.38	2.24	968	1.49	0.84	761	1.76	1.35
Onion												
Area	431	1,527	3,714	342	2,036	5,455	743	1,074	2,489	510	1,485	4,019
Output	431	13,885	29,608	342	17,539	36,355	743	14,587	51,185	510	24,189	99,249
Productivity	431	11.71	7.93	342	12.03	8.58	743	11.34	7.48	510	11.38	7.92
Potato												
Area	351	2,028	4,026	303	2,303	6,024	674	3,014	9,512	587	3,694	12,041
Output	351	28,503	44,128	303	27,843	43,051	674	67,058	248,196	587	71,627	286,377
Productivity	351	13.75	7.51	303	12.93	7.79	674	12.64	7.55	587	11.76	8.19
Rice												
Area	667	64,626	82,739	544	67,299	85,705	1017	88,839	104,258	784	100,968	120,405
Output	667	173,077	285,059	544	194,407	303,283	1017	160,160	221,919	784	197,829	266,243
Productivity	667	2.30	1.01	544	2.51	1.10	1017	1.61	0.87	784	1.81	0.94
Sesamum												
Area	573	3,245	6,935	460	2,790	4,742	908	4,826	11,359	749	5,919	15,535
Output	573	1,220	3,198	460	1,119	2,212	908	1,805	5,529	749	2,032	6,103
Productivity	573	0.35	0.23	460	0.38	0.25	908	0.32	0.22	749	0.35	0.24
Sugarcane												
Area	523	12,161	23,096	419	11,554	22,413	907	8,554	25,972	711	8,866	27,790
Output	523	955,008	1,797,426	419	902,855	1,738,094	907	590,206	1,786,733	711	588,924	1,878,506
Productivity	523	70.26	35.51	419	67.35	39.47	907	53.13	26.72	711	55.86	30.25
Tobacco												
Area	166	7,958	16,242	176	8,267	17,829	258	454	1,647	213	620	2,082
Output	166	9,853	22,353	176	10,113	20,766	258	663	2,233	213	1,128	3,622
Productivity	166	1.54	1.53	176	1.53	1.61	258	1.63	1.88	213	1.71	1.57
Wheat												
Area	437	60,088	81,807	349	64,550	81,240	923	49,803	65,451	689	52,869	67,471
Output	437	204,344	353,065	349	225,183	353,261	923	126,363	200,516	689	147,671	224,604
Productivity	437	2.21	1.25	349	2.38	1.27	923	1.78	0.97	689	1.93	1.02

Source: Rainfall data from TRMM satellite, crop data from State Agricultural Reports. Sample includes years 2000-2010 for districts falling within 5 thousand persons per branch of the national average. Observations are crop-years; the number of districts varies by crop as not every crop is grown in all districts. 302 of 572 districts are eligible for sample. Area is reported in Hectares square, output in tonnes, and productivity is output divided by area. Cotton reported in bales instead of tonnes.

Annual survey of industries

	Banked, Pre-reform			Banked, Post-reform			Under Banked, Pre-reform			Under Banked, Post-reform		
	N	Mean	Std. Dev.	N	Mean	Std. Dev.	N	Mean	Std. Dev.	N	Mean	Std. Dev.
Log Total Employees	42702	3.786	1.403	40252	3.954	1.436	21133	3.567	1.345	17976	3.72	1.403
Log Number of units	42824	0.04	0.193	40575	0.041	0.203	21216	0.021	0.15	18123	0.025	0.159
Plant Age	42248	16.002	13.986	39268	15.204	13.878	20864	14.97	14.197	17562	14.664	14.332
Log Capital (No Land or Inventory)	42339	14.911	2.876	39707	15.151	3.392	21030	14.576	2.952	17886	14.995	3.135
Log Net Assets	42352	15.679	2.883	39772	15.76	3.294	21040	15.354	2.929	17902	15.602	3.024
Log Working Capital	35823	15.306	3.024	34057	15.259	3.689	18262	15.015	3.105	15818	15.287	3.154
Log Loans	34828	14.869	4.037	32543	14.962	4.199	16258	14.874	4.084	13795	15.062	4.035
Log Total Investment	39950	14.688	3.2	37858	14.943	3.829	20517	14.248	3.298	17468	14.649	3.619
Capital Labor Ratio	42221	6.644	47.52	39543	11.121	237.379	20971	8.133	38.898	17800	10.879	105.471
Log Capital Labor Ratio	42202	0.774	1.535	39535	0.875	1.516	20958	0.89	1.662	17798	1.003	1.645

Source: Annual Survey of Industries, Unit level data 1999-2010. Sample is restricted to plants reporting being open and reporting a valid urban or rural status. Captital Labor Ratio constructed as average of opening and closing Net Assets divided by the total wage bill plus benefits. States and UTs selected by their share of population being concentrated on one side of the threshold or the other. "Banked States" include Haryana, Uttarakhand, Punjab, Mizoram, Daman and Dimiu, Karnataka, Puducherry, and "Under Banked States" include Rajasthan, Tripura, Jharkhand, Orissa, Dadra and Nagar Haveli.

Source: Annual Survey of Industries, Ministry of Agriculture, RBI and author's calculations.

Table 4: Fuzzy RD: private bank branches

Licences

VARIABLES	(1) 2002	(2) 2003	(3) 2004	(4) 2005	(5) 2006	(6) 2007	(7) 2008	(8) 2009	(9) 2010	(10) 2011	(11) 2012
Conventional	-0.0163	0.279	0.463	0.609	0.760	2.387**	4.086***	6.074***	7.793***	9.085***	10.26***
	[0.180]	[0.312]	[0.522]	[0.659]	[0.872]	[1.093]	[1.396]	[1.915]	[2.438]	[2.784]	[3.023]
Bias-corrected	-0.0500	0.356	0.928*	1.027	1.190	2.851***	4.509***	6.715***	8.311***	9.586***	10.82***
	[0.180]	[0.312]	[0.522]	[0.659]	[0.872]	[1.093]	[1.396]	[1.915]	[2.438]	[2.784]	[3.023]
Robust	-0.0500	0.356	0.928	1.027	1.190	2.851**	4.509***	6.715***	8.311***	9.586***	10.82***
	[0.214]	[0.370]	[0.631]	[0.795]	[1.033]	[1.301]	[1.647]	[2.265]	[2.890]	[3.293]	[3.573]
Bandwidth	3.500	3.500	3.500	3.500	3.500	3.500	3.500	3.500	3.500	3.500	3.500
N_Banked	95	94	94	94	94	94	94	94	94	94	94
N_UBanked	122	122	122	122	122	122	122	122	122	122	122
DepMean	8.714	8.917	9.241	9.847	10.62	11.92	13.83	15.31	17.13	18.47	19.99

Operating branches

VARIABLES	(1) 2002	(2) 2003	(3) 2004	(4) 2005	(5) 2006	(6) 2007	(7) 2008	(8) 2009	(9) 2010	(10) 2011	(11) 2012
Conventional	0.181	0.343	0.577	0.644	0.719	1.270	3.262**	4.840***	7.051***	9.219***	10.58***
	[0.152]	[0.320]	[0.557]	[0.641]	[0.865]	[1.005]	[1.279]	[1.653]	[2.159]	[2.718]	[3.102]
Bias-corrected	0.166	0.383	1.036*	1.036	1.139	1.468	3.754***	5.158***	7.414***	9.730***	11.11***
	[0.152]	[0.320]	[0.557]	[0.641]	[0.865]	[1.005]	[1.279]	[1.653]	[2.159]	[2.718]	[3.102]
Robust	0.166	0.383	1.036	1.036	1.139	1.468	3.754**	5.158***	7.414***	9.730***	11.11***
	[0.180]	[0.376]	[0.665]	[0.771]	[1.027]	[1.179]	[1.507]	[1.956]	[2.558]	[3.214]	[3.660]
Bandwidth	3.500	3.500	3.500	3.500	3.500	3.500	3.500	3.500	3.500	3.500	3.500
N_Banked	95	94	94	94	94	94	94	94	94	94	94
N_UBanked	122	122	122	122	122	122	122	122	122	122	122
DepMean	8.636	8.801	9.125	9.597	10.34	10.87	12.25	14.42	16.19	17.91	20.00

Operating banks

VARIABLES	(1) 2002	(2) 2003	(3) 2004	(4) 2005	(5) 2006	(6) 2007	(7) 2008	(8) 2009	(9) 2010	(10) 2011	(11) 2012
Conventional	-0.0503	-0.0960	0.00734	-0.109	-0.0831	0.400	0.932	1.095*	1.232*	1.463**	1.169**
	[0.126]	[0.240]	[0.287]	[0.324]	[0.451]	[0.453]	[0.611]	[0.612]	[0.639]	[0.605]	[0.583]
Bias-corrected	-0.0931	-0.0851	0.0320	-0.0943	0.00725	0.446	1.043*	1.131*	1.296**	1.557**	1.224**
	[0.126]	[0.240]	[0.287]	[0.324]	[0.451]	[0.453]	[0.611]	[0.612]	[0.639]	[0.605]	[0.583]
Robust	-0.0931	-0.0851	0.0320	-0.0943	0.00725	0.446	1.043	1.131	1.296*	1.557**	1.224*
	[0.149]	[0.279]	[0.338]	[0.388]	[0.540]	[0.540]	[0.714]	[0.716]	[0.755]	[0.714]	[0.683]
Bandwidth	3.500	3.500	3.500	3.500	3.500	3.500	3.500	3.500	3.500	3.500	3.500
N_Banked	95	94	94	94	94	94	94	94	94	94	94
N_UBanked	122	122	122	122	122	122	122	122	122	122	122
DepMean	2.788	2.963	3.171	3.532	3.991	4.269	4.787	5.421	5.875	5.856	6.250

Source: Population Census 2001, RBI, TRMM and author's calculations.

Note: Standard errors in brackets. ***p<0.01, **p<0.05,*p<0.1 for all tables. Estimated using local linear regressions with controls for district population and its square, the per cent of workers in agriculture, a control for monsoon rainfall and the pre-randomisation 2001 value of the dependent variable. Bandwidths are set 3,500 persons per branch and estimated using a triangular kernel. under-banked status is instrumented for with predicted under-banked assignment. Licences are considered in operation if they are granted for a branch currently operating or pending opening.

Table 5: Fuzzy RD: Private sector banks credit

AGGREGATE

Private sector credit: accounts

VARIABLES	(1) 2002	(2) 2003	(3) 2004	(4) 2005	(5) 2006	(6) 2007	(7) 2008	(8) 2009	(9) 2010	(10) 2011
Conventional	-0.186	0.386	1.695	3.652**	6.220*	10.57***	8.679*	8.432*	13.16**	18.54*
	[0.599]	[0.804]	[1.123]	[1.528]	[3.426]	[3.399]	[5.098]	[4.521]	[6.219]	[10.26]
Bias-corrected	-0.209	0.797	2.539**	4.172***	7.271**	12.18***	9.761*	9.950**	14.82**	19.66*
	[0.599]	[0.804]	[1.123]	[1.528]	[3.426]	[3.399]	[5.098]	[4.521]	[6.219]	[10.26]
Robust	-0.209	0.797	2.539*	4.172**	7.271*	12.18***	9.761	9.950*	14.82**	19.66*
	[0.698]	[0.955]	[1.307]	[1.835]	[4.047]	[4.228]	[5.988]	[5.302]	[7.279]	[11.92]
Bandwidth	3.500	3.500	3.500	3.500	3.500	3.500	3.500	3.500	3.500	3.500
N_Banked	95	94	94	94	94	94	92	94	94	94
N_UBanked	122	122	121	122	122	122	122	122	122	122
DepMean	5.067	5.484	6.470	8.800	12.83	13.77	16.14	17.78	22.82	25.80

Private sector credit: amounts

VARIABLES	(1) 2002	(2) 2003	(3) 2004	(4) 2005	(5) 2006	(6) 2007	(7) 2008	(8) 2009	(9) 2010	(10) 2011
Conventional	-484.3	-692.6	-9.290	1,096	1,442	2,187	2,163	1,538	1,580	2,288
	[1,524]	[1,984]	[1,912]	[1,305]	[1,680]	[2,135]	[2,031]	[1,412]	[1,287]	[1,568]
Bias-corrected	-31.48	-111.9	502.1	2,007	2,441	3,241	2,763	2,765*	2,403*	3,121**
	[1,524]	[1,984]	[1,912]	[1,305]	[1,680]	[2,135]	[2,031]	[1,412]	[1,287]	[1,568]
Robust	-31.48	-111.9	502.1	2,007	2,441	3,241	2,763	2,765	2,403	3,121
	[1,750]	[2,292]	[2,193]	[1,572]	[2,075]	[2,607]	[2,515]	[1,819]	[1,598]	[1,972]
Bandwidth	3.500	3.500	3.500	3.500	3.500	3.500	3.500	3.500	3.500	3.500
N_Banked	95	94	94	94	94	94	92	94	94	94
N_UBanked	122	122	121	122	122	122	122	122	122	122
DepMean	2641	3223	2943	3466	3922	4920	5278	5362	4932	5990

DIRECT AGRICULTURE AND PERSONAL LOANS

Private credit to direct agriculture and personal loans: accounts

VARIABLES	(1) 2002	(2) 2003	(3) 2004	(4) 2005	(5) 2006	(6) 2007	(7) 2008	(8) 2009	(9) 2010	(10) 2011
Conventional	-0.0122	0.812	1.742*	4.144***	6.899**	9.699***	6.494*	8.704***	13.55***	18.41**
	[0.386]	[0.569]	[0.961]	[1.357]	[3.125]	[2.590]	[3.602]	[3.342]	[4.557]	[8.812]
Bias-corrected	0.144	1.287**	2.406**	4.617***	7.781**	11.70***	7.807**	10.35***	15.55***	19.94**
	[0.386]	[0.569]	[0.961]	[1.357]	[3.125]	[2.590]	[3.602]	[3.342]	[4.557]	[8.812]
Robust	0.144	1.287*	2.406**	4.617***	7.781**	11.70***	7.807*	10.35***	15.55***	19.94*
	[0.460]	[0.665]	[1.112]	[1.647]	[3.725]	[3.330]	[4.317]	[3.963]	[5.335]	[10.21]
Bandwidth	3.500	3.500	3.500	3.500	3.500	3.500	3.500	3.500	3.500	3.500
N_Banked	95	94	94	94	94	94	92	94	94	94
N_UBanked	122	122	121	122	122	122	122	122	122	122
DepMean	3.311	3.460	4.640	6.357	9.937	9.638	10.82	13.69	15.78	17.77

Private credit to direct agriculture and personal loans: amount

VARIABLES	(1) 2002	(2) 2003	(3) 2004	(4) 2005	(5) 2006	(6) 2007	(7) 2008	(8) 2009	(9) 2010	(10) 2011
Conventional	43.37	85.10	771.3*	1,542**	1,792**	2,246***	1,381	930.4	1,121***	1,393**
	[45.30]	[60.71]	[465.7]	[633.3]	[712.2]	[865.0]	[933.8]	[595.4]	[409.0]	[581.7]
Bias-corrected	64.44	128.6**	900.6*	1,715***	2,088***	2,572***	1,287	1,014*	1,271***	1,538***
	[45.30]	[60.71]	[465.7]	[633.3]	[712.2]	[865.0]	[933.8]	[595.4]	[409.0]	[581.7]
Robust	64.44	128.6*	900.6*	1,715**	2,088**	2,572**	1,287	1,014	1,271***	1,538**
	[53.21]	[70.78]	[480.7]	[741.2]	[854.8]	[1,044]	[1,150]	[714.0]	[485.1]	[691.8]
Bandwidth	3.500	3.500	3.500	3.500	3.500	3.500	3.500	3.500	3.500	3.500
N_Banked	95	94	94	94	94	94	92	94	94	94
N_UBanked	122	122	121	122	122	122	122	122	122	122
DepMean	236.9	309.6	777.2	1209	1603	1978	1989	2121	1856	2269

Source: Population Census 2001, RBI, TRMM and author's calculations.

Note: Accounts reported in thousands. Amounts reported in millions of rupees. Estimated using local linear regressions with controls for district population and its square, the per cent of workers in agriculture, a control for monsoon rainfall and the pre-randomisation 2001 value of the dependent variable. Bandwidths are set 3,500 persons per branch and estimated using a triangular kernel. under-banked status is instrumented for with predicted under-banked assignment.

Table 6: RD from reduced form: credit from public sector banks

	Public sector credit accounts									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
VARIABLES	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
Conventional	-2.701	0.533	1.395	-2.342	-6.312	0.206	0.107	3.237	7.394	0.275
	[3.984]	[4.665]	[6.724]	[8.643]	[11.39]	[12.14]	[13.43]	[14.43]	[16.30]	[16.92]
Bias-corrected	-2.850	0.829	3.326	0.670	-5.159	4.182	3.492	9.762	11.80	2.916
	[3.984]	[4.665]	[6.724]	[8.643]	[11.39]	[12.14]	[13.43]	[14.43]	[16.30]	[16.92]
Robust	-2.850	0.829	3.326	0.670	-5.159	4.182	3.492	9.762	11.80	2.916
	[4.831]	[5.703]	[8.269]	[10.65]	[13.85]	[14.83]	[16.37]	[17.59]	[19.42]	[20.43]
Bandwidth	3.500	3.500	3.500	3.500	3.500	3.500	3.500	3.500	3.500	3.500
N_Banked	95	94	94	94	94	94	92	94	94	94
N_UBanked	122	122	121	122	122	122	122	122	122	122
DepMean	99.51	102.5	104.7	120.6	132.4	141.5	150.5	154.2	167.2	177.1

	Public sector credit amounts									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
VARIABLES	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
Conventional	1,156	847.0	792.4	319.8	-49.55	714.7	-161.0	3,247	3,371	4,283
	[754.8]	[925.1]	[1,195]	[2,949]	[3,985]	[3,628]	[5,479]	[4,186]	[4,281]	[4,860]
Bias-corrected	1,278*	1,214	1,118	785.4	481.1	1,953	1,685	5,274	4,712	5,939
	[754.8]	[925.1]	[1,195]	[2,949]	[3,985]	[3,628]	[5,479]	[4,186]	[4,281]	[4,860]
Robust	1,278	1,214	1,118	785.4	481.1	1,953	1,685	5,274	4,712	5,939
	[902.6]	[1,266]	[1,626]	[3,534]	[4,742]	[4,252]	[6,365]	[4,986]	[5,292]	[6,135]
Bandwidth	3.500	3.500	3.500	3.500	3.500	3.500	3.500	3.500	3.500	3.500
N_Banked	95	94	94	94	94	94	92	94	94	94
N_UBanked	122	122	121	122	122	122	122	122	122	122
DepMean	10586	11953	13479	17693	21386	23326	27222	29581	31372	34125

Source: Population Census 2001, RBI, TRMM and author's calculations.

Note: Accounts reported in thousands. Amounts reported in millions of rupees. Estimated using local linear regressions with controls for district population and its square, the per cent of workers in agriculture, a control for monsoon rainfall and the pre-randomisation 2001 value of the dependent variable. Bandwidths are set 3,500 persons per branch and estimated using a triangular kernel. Public sector banks include the State Bank of India and Associated Banks, Nationalised Banks, IDBI and Regional Rural Banks.

Table 7: Fuzzy RD: percentage change in private credit amount to rural and semi-urban areas

Direct to agriculture

VARIABLES	(1) 2002	(2) 2003	(3) 2004	(4) 2005	(5) 2006	(6) 2007	(7) 2008	(8) 2009	(9) 2010	(10) 2011
Conventional	0.0552	0.143	0.216	0.709**	1.029**	0.830*	0.908	1.840***	1.445**	1.755***
	[0.107]	[0.166]	[0.212]	[0.330]	[0.414]	[0.503]	[0.603]	[0.620]	[0.647]	[0.613]
Bias-corrected	0.0866	0.253	0.298	0.931***	1.273***	1.038**	1.195**	2.103***	1.713***	1.923***
	[0.107]	[0.166]	[0.212]	[0.330]	[0.414]	[0.503]	[0.603]	[0.620]	[0.647]	[0.613]
Robust	0.0866	0.253	0.298	0.931**	1.273**	1.038*	1.195*	2.103***	1.713**	1.923***
	[0.131]	[0.197]	[0.247]	[0.395]	[0.504]	[0.617]	[0.722]	[0.751]	[0.791]	[0.739]
Bandwidth	3.500	3.500	3.500	3.500	3.500	3.500	3.500	3.500	3.500	3.500
N_Banked	95	94	94	94	94	94	92	94	94	94
N_UBanked	122	122	121	122	122	122	122	122	122	122
DepMean	-0.0700	0.0481	0.164	0.433	0.550	0.964	1.488	1.419	1.953	2.376

Indirect to agriculture

VARIABLES	(1) 2002	(2) 2003	(3) 2004	(4) 2005	(5) 2006	(6) 2007	(7) 2008	(8) 2009	(9) 2010	(10) 2011
Conventional	0.0371	0.373	0.573	0.529	0.578	-0.0714	0.256	1.660**	1.724***	2.197***
	[0.268]	[0.355]	[0.426]	[0.565]	[0.491]	[0.609]	[0.709]	[0.682]	[0.539]	[0.595]
Bias-corrected	0.127	0.489	0.710*	0.710	0.812*	0.0273	0.316	1.992***	2.051***	2.493***
	[0.268]	[0.355]	[0.426]	[0.565]	[0.491]	[0.609]	[0.709]	[0.682]	[0.539]	[0.595]
Robust	0.127	0.489	0.710	0.710	0.812	0.0273	0.316	1.992**	2.051***	2.493***
	[0.306]	[0.432]	[0.516]	[0.678]	[0.591]	[0.733]	[0.871]	[0.820]	[0.646]	[0.714]
Bandwidth	3.500	3.500	3.500	3.500	3.500	3.500	3.500	3.500	3.500	3.500
N_Banked	95	94	94	94	94	94	92	94	94	94
N_UBanked	122	122	121	122	122	122	122	122	122	122
DepMean	0.317	0.199	0.188	0.237	0.257	0.453	0.955	1.039	1.313	1.133

Source: Population Census 2001, RBI, TRMM and author's calculations.

Note: Percentage change is approximated using difference in logs relative the value reported in 2001. Estimated using local linear regressions with controls for district population and its square, the per cent of workers in agriculture and a control for monsoon rainfall. Bandwidths are set 3,500 persons per branch and estimated using a triangular kernel.

Table 8: RD Results: crop yield index

Fuzzy RD Estimated annually, instrumenting for under-banked status

VARIABLES	(1) 2002	(2) 2003	(3) 2004	(4) 2005	(5) 2006	(6) 2007	(7) 2008	(8) 2009	(9) 2010
Conventional	0.105	-0.0327	-0.0324	0.250	0.160	0.158	0.0846	0.353*	0.251
	[0.162]	[0.205]	[0.194]	[0.170]	[0.194]	[0.266]	[0.225]	[0.204]	[0.191]
Bias-corrected	0.136	-0.0366	-0.00198	0.327*	0.217	0.178	0.0983	0.426**	0.317*
	[0.162]	[0.205]	[0.194]	[0.170]	[0.194]	[0.266]	[0.225]	[0.204]	[0.191]
Robust	0.136	-0.0366	-0.00198	0.327	0.217	0.178	0.0983	0.426*	0.317
	[0.200]	[0.254]	[0.243]	[0.206]	[0.240]	[0.325]	[0.269]	[0.256]	[0.232]
Bandwidth	3.500	3.500	3.500	3.500	3.500	3.500	3.500	3.500	3.500
N_Banked	77	77	78	79	74	75	78	74	70
N_UBanked	108	104	103	108	104	93	106	102	87
DepMean	0.121	-0.0216	0.106	0.117	0.119	0.0945	0.107	0.0857	0.0860

Fuzzy RD instrumenting for private bank credit accounts, pre-reform and post-reform

VARIABLES	(1) preref	(2) postref
Conventional	-0.0412	0.0230*
	[0.168]	[0.0138]
Bias-corrected	-0.0417	0.0264*
	[0.168]	[0.0138]
Robust	-0.0417	0.0264
	[0.206]	[0.0175]
Bandwidth	3.500	3.500
N_Banked	230	442
N_UBanked	314	600
DepMean	0.0685	0.102

Standard errors in brackets

*** p<0.01, ** p<0.05, * p<0.1

Source: Ministry of Agriculture, Population Census 2001, RBI, TRMM and author's calculations.

Note: Index of crop yield using weighted averages of the crops rice, wheat, jowar, groundnut and cotton. Weighted by crop revenue share. Estimated using local linear regressions with controls for district population and its mean, the per cent of workers in agriculture, a control for monsoon rainfall, and year fixed effects. No pre-randomisation value of the dependent variable is included. Bandwidths are set at 3,500 persons per branch and estimated using a triangular kernel. Pre-reform years are considered 2002-04 and post-reform is 2005-10.

Table 9: Fuzzy RD: percentage change in private credit amount to manufacturing and processing from 2001 level

VARIABLES	(1) 2002	(2) 2003	(3) 2004	(4) 2005	(5) 2006	(6) 2007	(7) 2008	(8) 2009	(9) 2010	(10) 2011
Conventional	-0.145	0.0706	0.127	-0.331	-0.587	-0.00749	0.522	1.065*	1.437**	0.991
	[0.489]	[0.585]	[0.524]	[0.486]	[0.633]	[0.579]	[0.612]	[0.618]	[0.670]	[0.667]
Bias-corrected	-0.143	0.170	0.183	-0.280	-0.624	-0.0967	0.662	1.421**	1.824***	1.295*
	[0.489]	[0.585]	[0.524]	[0.486]	[0.633]	[0.579]	[0.612]	[0.618]	[0.670]	[0.667]
Robust	-0.143	0.170	0.183	-0.280	-0.624	-0.0967	0.662	1.421*	1.824**	1.295
	[0.595]	[0.717]	[0.645]	[0.613]	[0.778]	[0.696]	[0.736]	[0.753]	[0.817]	[0.805]
Bandwidth	3.500	3.500	3.500	3.500	3.500	3.500	3.500	3.500	3.500	3.500
N_Banked	95	94	94	94	94	94	92	94	94	94
N_UBanked	122	122	121	122	122	122	122	122	122	122
DepMean	0.934	1.098	0.678	0.763	0.553	0.694	1.119	1.231	1.287	1.410

Source: Population Census 2001, RBI, TRMM and author's calculations.

Note: Percentage change is approximated using difference in logs relative the value reported in 2001. Estimated using local linear regressions with controls for district population and its square, the per cent of workers in agriculture, and a control for monsoon rainfall. Bandwidths are set 3,500 persons per branch and estimated using a triangular kernel.

Table 10: Diff n Diff: States selected around under-banked threshold, 1999-2010

VARIABLES	(1) Ln_Net_Assets	(2) Ln_Working_Capital	(3) Ln_Loans	(4) Ln_Tot_Investment	(5) Cap_Labor_Ratio
TreatXPost2006	0.171	0.264*	0.235*	0.197*	3.426*
	[0.142]	[0.136]	[0.116]	[0.106]	[1.724]
Observations	118,236	101,566	95,269	113,296	118,128
R-squared	0.270	0.195	0.082	0.200	0.012
State FEs	Yes	Yes	Yes	Yes	Yes
State Trend	Yes	Yes	Yes	Yes	Yes

Robust standard errors in brackets
*** p<0.01, ** p<0.05, * p<0.1
Standard Errors Clustered at State level

Source: Annual Survey of Industries and author's calculations.

Note: Banked states include Haryana, Uttarakhand, Punjab, Mizoram, Daman and Dimiu, Karnataka and Puducherry. Under-banked states include Rajasthan, Tripura, Jharkhand, Orissa and Dadra and Nagar Haveli. All regressions control for post-2006 and treated state individual effects, logged number of units in firm and the logged number of employees in the enterprise, plant age and its square, a year trend, state specific year trends and state fixed effects.

Table 11: Difference in log mean district light from 2004

Fuzzy RD estimated annually, instrumenting for under-banked status

VARIABLES	(1) 2005	(2) 2006	(3) 2007	(4) 2008	(5) 2009	(6) 2010	(7) 2011	(8) 2012
Conventional	-0.0373	0.00436	0.113***	0.112**	0.0479	0.126	0.111*	0.119
	[0.0227]	[0.0325]	[0.0415]	[0.0460]	[0.0778]	[0.0773]	[0.0652]	[0.105]
Bias-corrected	-0.0430*	0.00773	0.129***	0.128***	0.0567	0.151*	0.141**	0.148
	[0.0227]	[0.0325]	[0.0415]	[0.0460]	[0.0778]	[0.0773]	[0.0652]	[0.105]
Robust	-0.0430*	0.00773	0.129***	0.128**	0.0567	0.151*	0.141**	0.148
	[0.0244]	[0.0358]	[0.0445]	[0.0501]	[0.0849]	[0.0845]	[0.0700]	[0.114]
Bandwidth	3.500	3.500	3.500	3.500	3.500	3.500	3.500	3.500
N_Banked	95	95	95	95	95	95	95	95
N_UBanked	122	122	122	122	122	122	122	122
DepMean	-0.139	-0.0805	0.114	0.0722	0.0266	0.355	0.219	0.297

Fuzzy RD instrumenting for private bank branches, pre-reform and post-reform

VARIABLES	(1) preref	(2) postref
Conventional	-0.0264	0.00508
	[0.274]	[0.00373]
Bias-corrected	-0.0455	0.0115***
	[0.274]	[0.00373]
Robust	-0.0455	0.0115***
	[0.320]	[0.00444]
Bandwidth	3.500	3.500
N_Banked	94	658
N_UBanked	122	854
DepMean	-0.139	0.143

Standard errors in brackets

*** p<0.01, ** p<0.05, * p<0.1

Source: NOAA, Population Census 2001, RBI, TRMM and author's calculations.

Note: Reduced form estimated using local quadratic regressions with controls for district population and its square. Bandwidths are set to 3,500 persons per branch and estimated using a triangular kernel. The fuzzy regression discontinuity is estimated using local linear regressions. The number of operating private bank branches is instrumented with predicted under-banked assignment. Controls include district population and its square, the per cent of workers in agriculture, and a control for monsoon rainfall. Pre-reform year is 2005 using 2004 as the base year for the approximate percentage change. Post-reform years are 2006-12.

Table 12: NREGA discontinuity in district phase assignment

	(1)	(2)	(3)
VARIABLES	Phase_1	Phase_2	Phase_3
Conventional	-0.0648	0.0145	0.0503
	[0.119]	[0.0909]	[0.135]
Bias-Corrected	-0.121	0.0710	0.0497
	[0.119]	[0.0909]	[0.135]
Robust	-0.121	0.0710	0.0497
	[0.139]	[0.109]	[0.160]
Bandwidth	3.500	3.500	3.500
N_Banked	93	93	93
N_UBanked	121	121	121
DepMean	0.285	0.201	0.514

Standard errors in brackets

*** p<0.01, ** p<0.05, * p<0.1

Source: NREGA, Population Census 2001, RBI and author's calculations.

Note: Reduced form estimated using local quadratic regressions with controls for district population and its square. Bandwidths are set to 3,500 persons per branch and estimated using a triangular kernel. NREGA was rolled out in three phases between 2006 and 2009 based on some measure of expected programme need by district.

Data appendix

The details of the data directly relevant to the analysis are discussed below. Additional descriptions of the data and their preparation, covering the harmonisation of district data for the panel, banking data on branches and credit, population group definitions, agricultural and industrial data, and remote sensing data are available in the online appendix to this paper. Online appendix: http://people.bu.edu/nvcyoung/Web_Appendix_FBEG.pdf

7.1 Constructing the forcing variable

In constructing the forcing variable and national average I follow the APPBO procedure [37] described for identifying deficit districts during the policies of the 1980s and also that for identifying under-banked states in the RBI Report of the Group to Review Branch Authorisation Policy (RBI, 2009). I take the number of operating branches on 7 September 2005, the day prior to the 2005 Master Circular issue date that implemented the branching policy reform. Following the rule that under-banked Status = 1(district population per branch $>$ national average) yields nearly an exact match to the official 2006 list of under-banked districts in the 2006 master circular.[38] Out of 572 districts only five fail to conform to their official status. Most are close to the cutoff, while the APPBO of one district places it outside the local linear regression bandwidth. Due to redistricting and the level of aggregation of credit and deposits data, I aggregate all districts bifurcating since 2001 back to their 2001 boundaries. In cases that new districts are formed from two or more source districts, these are aggregated into a single super district, resulting in 572 districts. Of these, I denote 202 districts as banked (with 203 on the official list) and 370 under-banked (369 officially). Super districts are dropped throughout the analysis. Replicating the analysis by taking the number of operating branches on 1 January 2006 yields similar results.

7.2 Crop yield index

Annual crop yield is calculated as crop output in tonnes per hectare cultivated for that crop. To create the index of crop yields as in Jayachandran (2006), I match the crop prices data to the crop output and area data. Four of the top five revenue producing crops for India identified in Jayachandran (2006) are used in the index: rice, wheat, jowar and groundnut. Cotton is substituted for sugar in the index, due to concerns regarding the accuracy of conversions of sugarcane to raw

[37]The Average Population Per Bank Office was constructed using the district population from the most recent population census, in this case that from 2001, and dividing that by the number of bank offices in that district. I restrict the set of offices to those conducting general and specialised bank business which may depart from the actual algorithm used by the RBI. The national average with which the value is compared is the total population of India divided by the number of bank offices.

[38]A list of under-banked districts was issued with the 2005 master circular as well. A slightly revised list was reissued with the 2006 master circular and remained unchanged through 2009, after which the districts of some states were dropped. The national average computed using 7 September 2005 as the policy date was 14,915 persons per branch in India.

sugar production in order to match the two datasets, and whether the reported prices for sugar capture actual prices faced by farmers after accounting for delay of payments bargaining. Crop yields are normalised to have mean values equal to one in each year for comparability across crops. Weighted averages of the log values of the four crop yields are taken for each district year, using the crop revenue share of the total crop revenue of the district from those four crops as weights. When matching the price and production datasets, season and variety matches are made when the detail of data from both sets allow. Otherwise, the mean of price data by district and crop are calculated (if price is broken out by variety or season) and matched to the production data for that crop year. To increase the number of matches, when prices are missing for a crop at the district level, the weighted state average prices provided in the reports are used. Missing crop prices at the district level generally correspond to relatively low levels of output in the production data. An index exclusively using weighted state average prices is also constructed. The index is currently constructed for 2002-10.