Banking and Growth:
Evidence from a Regression Discontinuity Analysis

Nathaniel Young*

January 27, 2019

Abstract

This paper examines the impact of banking expansion on economic growth utilizing a previously unstudied policy reform. Contrary to evidence from US and cross-country analysis, several microeconomic studies do not find enduring banking effects in developing countries. I exploit exogenous bank branch expansion in India, iterating a regression discontinuity design across years. Applying the strategy to banking and productive sector outcomes, I trace out effects consistent with the 2005 reform timing and incentives. I find strong causal evidence that the expansion of financial intermediation led to positive outcomes in agriculture and manufacturing, and confirm local GDP growth using nightlights data. JEL: G21, G28, L13, O12, O16

*European Bank for Reconstruction and Development. Email: youngna@ebrd.com. 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 advisors 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, Ralph de Haas, 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, NEUDC and Penn State 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.

1 Introduction

Strong financial systems contribute to the growth of an economy (King and Levine, 1993; Jayaratne and Strahan, 1996; Rajan and Zingales, 1998). Empirical studies of bank branch expansion in the United States during the early 20th century (Dehejia and Lleras-Muney, 2007) and interstate branching reforms from the 1990s (Krishnan et al., 2014) find strong positive effects in growth and firm productivity, respectively, from improved banking access. Yet, a puzzle has emerged in developing countries. Recent micro-level studies don't find lasting positive results from financial inclusion interventions, raising the question of whether financial access actually promotes development. Microfinance, once the leading idea in development finance, faces mixed empirical evidence and struggles to produce long term income growth for recipients (Karlan and Zinman, 2011; Banerjee et al., 2015). These issues are not restricted to microfinance (Fulford, 2013). Despite the mixed evidence, improving access to formal banking remains a popular intervention in developing economies and disadvantaged communities of richer countries.

I analyze a previously unstudied policy reform in India introduced in 2005, finding new and rigorous evidence of positive effects from bank branch expansion on economic growth. I make two significant contributions to the literature. First, I develop a robust identification strategy for analyzing the impact of expanded banking access and apply it to a wide range of outcomes. The geographic and temporal dimensions of the reform, utilized by iterating a regression discontinuity design, builds a solid foundation for making causal inferences. Second, my analysis provides strong causal evidence that bank branch expansion led to positive effects on agricultural productivity, manufacturing outcomes and local income growth. These results fill a gap in the literature opened by rececent critiques (Panagariya, 2006; Kochar, 2011) of the seminal findings in Burgess and Pande (2005).[1]

[1]Kochar (2011) and Panagariya (2006) review the timing and details of branching policies implemented around social banking, highlighting conflicts with the trend-break analysis underpinning the identification strategy in Burgess and Pande (2005), and suggest competing interventions in explaining those results.

Development finance struggles to successfully induce investment in productive assets (Kaboski and Townsend, 2012; Field et al., 2013; Angelucci et al., 2015; Banerjee et al., 2015), hindering long term growth. The seminal work of Burgess and Pande (2005) finds that rural banks opened during India's Social Banking period of the 1970s-80s reduced the incidence of poverty at state levels and raised rural agricultural wages. Fulford (2013) suggests the gains from Social Banking were short lived, providing evidence that initially expanding consumption levels contracted in the longer term. Credit availability increases current consumption and replaces cautionary savings in his model, leading to lower investment and growth. Kaboski and Townsend (2011, 2012) find similar contractions from microfinance in Thailand.

The positive effects on income growth found in this paper complement the findings described above. I document credit expansion to productive sectors and chart corresponding effects in the real economy as channels for growth. Private sector commercial banks account for this credit expansion, unlike the microfinance and government institutions driving the interventions studied in earlier contributions. I examine this nuance further using a theoretical framework that captures why these differences in results should be expected.

My identification strategy leverages geographic and temporal variation to isolate the causal effects of expanded banking access on economic growth. I exploit the rule used by regulators when they selected 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," emits a regression discontinuity design. The national average constitutes a threshold, where districts with higher populations per branch receive treatment and the others do not. The reform encouraged additional branch entry in treated districts. The regression discontinuity allows me to overcome the classic endogeneity concern that bank branch expansion is selective on growth potential. Moreover, the unique selection rule helps in disentangling effects from contemporaneous interventions.

I trace the impact of the policy on a variety of outcomes through time, using pre-reform years as placebos. Despite constantly evolving values of the

underlying number of branches per capita, districts' official treatment status remained essentially unchanged across years. I can therefore estimate the average treatment effect of the reform through time by separately estimating the regression discontinuity for each year from 2002 to 2012.

I use several data sources to connect banking outcomes with impacts on the real economy. Sources include 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 to control for monsoon variation. Detailed bank branch data are collected separately from credit statistics by the RBI, providing a good cross check for these 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.

The first set of results confirm that the policy reform resulted in a significant expansion of bank branches in underserved areas. Importantly, this response is concentrated in the expansion of private sector banks. The cumulative effect of the reform is an average additional 10 private bank branches per district by the start of 2012. This effect is large: approximately 50% of the per district sample average around the threshold.

The strong response from private sector banking sheds light on the positive real economy effects that I uncover but were missing from microfinance and other analyses. Private sector banks in India emerged as a viable sector following the banking reforms in the early 1990s. Private sector banks at the time of the 2005 reform arguably look more like the financial intermediaries from economic theory, disciplined by markets into allocating capital toward productive investments, than do their microfinance and government bank counterparts.

Nationalised, public sector banks and the regional rural banks they sponsor, dominated branch expansion and banking activity under the earlier Social Banking period of the 1970s-80s. These banks can face internally misaligned incentives that impede their quality of financial intermediation. Several micro-level studies find evidence from public sector banks in India of under lending to productive firms, inertia in credit limits and little improvement in deliver-

ing development oriented lending (Banerjee and Duflo, 2001; Banerjee et al., 2004; Cole, 2009; Banerjee and Duflo, 2014). Looking across countries, higher incidence of government ownership in banking correlates positively with slower growth (La Porta et al., 2002).

To interpret the empirical results, I provide a light theoretical framework. The framework incorporates adverse selection and switching costs, which is particularly helpful in understanding the impact of intensified competition. Profit maximizing incentives induce strategic responses by banks in their entry and lending behavior. Banks expand credit early to lock in consumers who face positive switching costs. Details of the policy reform create staggered incentives for entry and lending that can be checked using the time panel data. I find evidence from lending consistent with predictions of anticipation effects in the framework. Public sector lending showed no response, consistent with lower sensitivity to market incentives.

Expanding credit exhibited geographic and sectoral diversity. Agricultural credit increased in rural and semi-urban areas of treated districts near the cutoff. Besides its importance in showing credit served an important productive sector,[2] the finding demonstrates that private sector banks lent in less populated areas and did not solely direct rural deposits to urban centers. Investments in productive sectors help offset the concern of later downturns in case consumption loans undermine saving as in Fulford (2013).[3]

I estimate that the expansion of formal banking had significant positive impacts on agricultural productivity. I apply the regression discontinuity design to a revenue weighted index of crop yields. Combining nine years of district level crop statistics and harvest prices, I construct a new data set to calculate the yield. I find positive effects consistent with the pattern of expanding credit, estimating that an increase of 1,000 private bank credit accounts in a district raises average crop yield by 2.3%. This effect is a little less than one third of the effect from a positive rainfall shock on yield found in Jayachandran (2006).

[2]56% of Indian workers in 2001 engaged in agricultural endeavors.

[3]Personal loans also expanded substantially, with the potential to improve productivity through investments in education or health.

These findings suggest banking provides productive benefits and does not only displace informal lending services. These results are in line with studies finding banking effects in farming via farm labor supply (Jayachandran, 2006), and in cropping decisions in recent work by Allen and Atkin (2015). This paper abstracts from specific mechanisms driving efficiency gains.

Turning to manufacturing, I estimate positive effects on borrowing and capital. Using annual ASI data, I find that enterprises in states with populations most affected by the reform reported higher total investments, working capital and capital labor ratios. This analysis rests on a difference-in-differences empirical strategy, since the ASI data are at the state level, so results should be interpreted with appropriate caution.

The positive effects from productive sectors more closely align with findings from historical branch expansions in the United States than recent work on microfinance. Dehejia and Lleras-Muney (2007) find that increased U.S. bank branching from 1900-40 encouraged growth in agriculture and manufacturing. Increased branching activity following the Interstate Banking and Branching Act of 1994, examined by Krishnan et al. (2014), led to greater efficiency gains by previously credit constrained manufacturers. The impact on productive sectors found here perhaps eased initial resource misallocation a la Hsieh and Klenow (2009) with implications for TFP.

Indeed, I confirm aggregate effects on local GDP growth, finding higher rates of increased nighttime light intensity experienced in treated districts in the years following the reform. The nightlights data provide a reliable proxy for economic growth, overcoming 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 the total effect from branch expansion was an average increase of local GDP by 1.65% in treated districts. Overall, these findings offer strong causal evidence that the expansion of formal banking facilitates growth across productive sectors and encourages economic development.

The next section describes important aspects of India's banking system and branch licensing reforms utilized for this 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. Section 6 discusses results and section 7 concludes.

2 Policy Reform and Institutional Background

2.1

This paper utilizes a previously unstudied policy reform to bank branch licensing in India implemented on September 8, 2005. The banking sector in India does not permit free entry of banking firms or branches. Banks must acquire licenses from India's central bank, the Reserve Bank of India (RBI), prior to opening all new branches. Permissions are also required to close or shift branches in most markets.[4] Prior to the 2005 reform, banks applied for each change to their branch network on a case-by-case basis through regional offices of the RBI. Licensing policies did not broadly target markets with specific characteristics, such as requiring branch openings in rural areas, since the end of the Social Banking period in 1990.[5]

The reform in 2005 changed branch licensing in two fundamental ways. First, the total branch licenses issued to a bank was tied to their proposed number of openings in a set of districts the RBI designated as being under banked.[6] The rule adopted by regulators to select the set of under banked districts compared the average number of persons per branch in a district against a statistic termed the "national average" of population per branch for India (RBI, 2009). Though an exact quota of branches is not stated explicitly, I argue that a quota-like system required banks to expand in under banked

[4]Banks were not allowed to relocate branches if it left a market unbanked. Branches in under banked districts, described later in this section, could not relocate to banked ones. See the online appendix for additional details.

[5]The LEAD banking program and Service Area Approach were in operation, aimed at delivering banking development objectives. See the online appendix for details.

[6]Banks were also judged on their provision of "no-frills" accounts, meeting priority sector lending obligations and their handling of complaints. See the 2005 issue of the *Master Circular on Branch Authorisation*, issued annually by the RBI, for details of the reform.

districts in order to receive licenses for entry in rich markets. Most districts in India fall close to the national average in terms of population per branch. Those happening to fall above the cutoff received under banked status and the incentives for additional branch entry, despite looking otherwise similar to districts just below the cutoff. I discuss the empirical strategy based on this quasi-natural experiment in detail in section 4.

Second, the case-by-case application procedure for requesting new licenses was replaced with an Annual Branch Expansion Plan (ABEP) approach. Banks proposed a set of branch openings, closings and shifts to be implemented over one year. The RBI reviewed the list centrally, potentially met with bank management, and granted a set of permissions.[7]

Banks experienced far greater choice in selecting their locations for entry under the 2005 reform than under the policies of Social Banking. Unlike the 4:1 entitlement policy studied in Burgess and Pande (2005), which strictly required entry in unbanked markets, banks could choose among any of the markets in under banked districts to satisfy their obligated entry. In stark contrast to the planned approach to district-wise branch expansion implemented in the 1980s (RBI, 2009; Kochar, 2011), banks under the 2005 reform could choose their extensive and intensive level of entry in under banked districts, as well as the total expansion of their branch network. These differences enhanced the potential for direct competition between banks.

The banking environment in 2005 supported a much stronger and enabled private sector than under earlier reforms. 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 operate alongside and compete with government owned banks. The new private banks broadly face the same regulation as the other scheduled com-

[7]Permissions remained valid for one year, with the potential for extensions. Banks could submit their next ABEP early upon accomplishing 75% of their planned expansions.

mercial banks.[8] 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 Reform Details and Timing

The list of under banked districts remained nearly constant through the end of the sample period. Minor revisions to the 2005 list were introduced in 2006; the list was then reissued unchanged from 2007 to 2010.[9] 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 license issuance, I find lasting effects through 2012 as expected. Section 4 details how the empirical strategy exploits the persistence of under banked status in identifying the effect of banking on the real economy.

Further, heavy regulations governing the closing and shifting of branches limited subsequent adjustments of existing branches in a bank's network. Limiting network adjustments seems to have prevented extensive gaming in the system, such as opening or moving temporary branches across districts in order to satisfy entry requirements. Relatively few branch closures are observed in the data.[10]

In the online appendix, I discuss the opportunities for banks to behave strategically in the timing of their responses to the reform. The discussion

[8]Private sector banks carry the additional mandate of maintaining at least 25% of their branch network in population centers with fewer than one hundred thousand people.

[9]Starting in 2008, certain areas within under banked districts lost their status. State capitols, district headquarters and metropolitan areas no longer counted as under banked. Further, locations within 100 km of Mumbai, New Delhi, Kolkata and Chennai, and within 50 km of state capitols were made ineligible. Exceptions were made for Jammu and Kashmir, and the seven North Eastern states, Arunachal Pradesh, Assam, Manipur, Meghalaya, Mizoram, Nagaland and Tripura.

[10]During bank mergers, most branches of exiting banks are typically reopened under the acquiring bank, though sometimes as satellite offices.

covers grace periods, license validity length, the Vyas Committee and Service Area Approach reform.

2.3 Direct reform incentives

The 2005 reform purposefully introduced incentives for banks to make branching decisions between locations while taking their districts' under banked status into account. Licenses for branches in high profit areas could be used to leverage bank entry into locations of under banked districts. This mechanism requires high demand for branches in "rich" areas. The strong economic growth in India beginning in 2003 and continuing through the decade likely helped meet this requirement during the study period.

Banks could operate policy-driven branch openings in ways that minimized costs and performed very little banking activity besides accepting deposits. The branch licensing policies placed no branch-level requirements on the amount of banking activity. Individual branches must meet minimum staffing requirements, as well as minimum 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.

Individual branches are also not bound by Priority Sector lending ratios, important regulations that affect total bank lending.[11] These requirements must only be met at the bank level, and may be distributed unevenly across branches. In 2007, new guidelines reduced the set of loan categories eligible for priority sector status.[12] The reformed guidelines concentrated lending into direct and indirect agricultural endeavors, 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.

[11] Banks must maintain 40% of their outstanding credit in priority sector loans. Any shortfalls must be made up for with low interest lending to the NABARD RIDF fund. Banks typically come very close to meeting the requirement.

[12] 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

3.1

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 realized 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 subsidized required entry in lower performing districts with entry in richer ones. Finally, as the above mechanisms rely on incentives consistent with profit-maximizing objectives, different responses to the reform by private and public sector banks are predicted.

The framework adopts a standard characterization of financial intermediation with adverse selection of borrowers, a feature common to credit markets in developing economies.[13] 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 (normalized 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

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

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: 1) Banks possess a screening technology that reveals a potential borrower's type with certainty and costs amount s. 2) There exists a downward sloping demand curve among safe types. [14] 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.[15] 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 sizable

[14] 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.

[15] 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).

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.[16] 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.[17] 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.[18] 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 favorable expected

[16] 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.

[17] 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.

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

lending conditions offered in the future.

Assume banks are symmetric and profit maximizers, each facing an exogenous marginal cost of funds, including administrative costs from lending. Recall that banks cannot discriminate in the interest rate it offers to repeat versus first time borrowers. Since banks observe the parameters on the population defining the distribution of safe types, they know the slope of the demand curve, though 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 maximizes 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 maximizing 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 split the market.

However, if one bank acts as an incumbent, then it may choose to alter its behavior 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 2nd 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 maximizes 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 realized entry.[19]

3.2 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 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. Finally, the joint positive profits will be hardest to satisfy for policy eligible districts

[19] See the online appendix for additional discussion.

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.[20] 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.

3.3 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

[20]The strong Indian economy presumably drives branch openings in rich districts. Regulators could lever greater entry in under banked districts after the reform by increasing the licenses given in rich districts. Alternativley, setting under banked branching requirements too high could cool off that demand.

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 reform, and again upon realized 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 maximization 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

4.1

The rule adopted by the RBI in the 2005 reform to select the set of districts designated as under banked yields a clear quasi-natural experiment, exploitable by regression discontinuity. With geographic and temporal variation, the environment is nearly ideal for identifying the effects of banking access. The design applied here addresses the classic endogeneity concerns outlined in Burgess and Pande (2005) that traditionally frustrate causal inference.

The methodology assigning under banked districts used two inputs. First, the total 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 2005-2006 to obtain a "national average of popu-

lation per branch." The analogous value calculated for each district and set against this national average determined district status. Most districts have values close to the cutoff. Those districts with a local value higher than the national value received under banked status.

In terms of regression discontinuity design, the above algorithm induces a cutoff at the value of the national average, with district population per branch serving as the forcing variable. The policy generates an arbitrary difference in districts falling around the cutoff. Districts on the "under banked" side provide additional incentives for branch entry as described in section 2. Districts falling on the other side do not, despite being otherwise similar. Thus, the policy produces different probabilities of branch entry on either side of the cutoff. This regression discontinuity design is 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.

The McCrary test checks for perfect manipulation of treatment status by examining the running variable around the threshold for bunching (McCrary, 2008). I find that the distribution of districts is smooth at the threshold, and fail to reject the null hypothesis of continuity with the test showing a log difference in height of 6.6 and a standard error of 22. The lack of manipulation is not surprising due to the assignment rule's construction. The population census occurred in 2001, four years prior to the policy. District treatment status could then only be altered through changes to the number of operating branches. Any branching changes would be conditional on RBI approval, accrue across competing banks simultaneously and move the value of the cutoff in addition to any district's individual measure. Any manipulation strategies faced both uncertainty and a moving target.

The second validity check tests whether other factors that may affect the outcomes of interest are continuous in the running variable at the cutoff. District baseline values for population, the share of district population accounted for in its four largest cities, the scheduled caste and tribes population, the percent literate, percent working, percent of workers in agriculture, all taken from the 2001 Population Census, and the number of private sector operat-

ing branches in 2000, are considered as potential factors. These baselines are tested for discontinuity at the threshold by separately performing RD analysis, taking each factor as the dependent variable. I do not find evidence of discontinuities in the baseline characteristics, supporting the assertion of proper randomization of districts around the cutoff.[21]

Figure 1 presents the tests visually in a series of plots, 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 a flexible slope on either side of the cutoff. Beyond demonstrating the continuity of these factors, the figure reveals broader trends in branching behavior at the time of the policy reform. Districts left of the cutoff, with more branches per person by definition, also tended to be places with more highly concentrated populations, exhibited higher literacy rates, had lower underprivileged populations and had a lower share of workers engaged primarily in agriculture.[22]

The assignment rule produced substantial geographic diversity with a high density of branches near the cutoff. The district populations per branch mapped in panel A of figure 2 shows the variation in the running variable across India.[23] Panel B shows that treatment and control districts used in estimation, being near enough to the cutoff, are drawn evenly from across the country. Additional checks of the RD design can be found in the online appendix, including regional level analysis.

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 observa-

[21] Regression results are presented in the online appendix.

[22] District rainfall, important for agriculture, is smooth at the cutoff. See the online appendix.

[23] The districts with greater deficits of branches per person, denoted by darker colors, broadly match the under served areas identified in the 2004 Vyas Committee's map.

tions 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 and Lemieux, 2010). I restrict all analysis to local linear and local 2nd degree polynomial regressions as recommended in Gelman and Imbens (2018). I set the bandwidth at 3.5 thousand 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.[24] 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,

$$y_i = \alpha + D_i\tau + f(PopPerBranch - Cutoff) + \delta X_i + \epsilon_i \quad (1)$$

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-randomization 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

[24]Results are robust to different bandwidth selections. 2nd degree polynomials typically perform better with wider bandwidths than linear specifications, as in the example from Lee and Lemieux (2010).

well as the percent 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 realized 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 realized list.[25] 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 "residualize" the data, regressing y_i on the set of controls X_i from equation 1, then estimating equation 1 replacing the left hand side variable with the residuals obtained from the first regression and dropping 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

Connecting real economy outcomes to realized exogenous changes in banking outcomes, as opposed to treatment status, is directly interesting. For agricultural and income growth outcomes, I therefore estimate the effects with the fuzzy RD instrumenting for banking outcomes.[26] That estimate provides in-

[25]Five out of 572 districts fail to match their realized under banked status from the 2006 list. See the data appendix for details.

[26]I also present the reduced form effects of under banked status.

sight into the effect of the specific banking outcome on the economic outcome of interest. However, that effect should be interpreted with care, as reform status influences 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, and can be compared with predictions from the theoretical framework of section 3. To examine 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, i.e. nighttime light. Recall from section 2, the RBI's list of under banked districts remained essentially unchanged in the reform period. I therefore hold the forcing variable fixed across regressions. The annual estimates of the discontinuity captures the short and medium term policy effects as they emerge.

4.2 Manufacturing

To examine effects in the manufacturing sector, I use ASI data available at the state level. The level of aggregation prevents applying the regression discontinuity directly. Instead, I follow a difference in differences approach, utilizing the institutional knowledge of the reforms to construct sets of treatment and control states.

I select the set of "under banked treatment states" as follows. Using population census data at the district level, I construct the shares of state population in under banked districts. 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 thousand 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. I will cautiously take the estimate as suggestive of the effect from the policy reform on manufacturing.

5 Data

The primary data on banking are from data sets 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 is 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 is usually impossible. Fortunately, the policy reform applied to districts, allowing analysis directly at the level of the reform. Utilizing the time dimension further helps to disentangle effects of the reform.

The credit limits, amounts and accounts data reported to the RBI are delineated by their intended geographic area of utilization. The use of Call Reports from banks do not typically allow for this level of geographic precision in terms of the utilization of funds, distinguishing this analysis from other work. This feature increases confidence that we are measuring the local availability of credit.

To conduct the analysis on agriculture, I develop a new data set 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 - 2010. 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% sample of firms representative at the state level. The ASI data used in this analysis does 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 nighttime 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

Results detail the expansion of private sector bank branches in treated districts. An anticipatory expansion of credit occurs, while no response is detected from public sector banks, consistent with the theoretical framework. Positive effects are estimated for the real economy.

6.1 Banking

To motivate the primary set of empirical results, I present a before and after visual example from two years. Figure 3 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 on January 1st of the respective year, 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 centered 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.

To reduce the noise in the estimated discontinuity and show its dynamics around the reform, I estimate the reduced form equation 1 separately for each year, with operating private branches as the dependent variable.[27] I plot the intercept points at the cutoff from each annual local linear regression by year in the right panel of figure 4. 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% level.[28]

The policy effect clearly emerges after 2006. Steadily higher branch growth in under banked districts produces the expanding positive discontinuities in 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 randomization at the cutoff. The muted response in 2006 and 2007 is consistent with most banks utilizing the policy grace period, waiting to submit their first ABEP until mid 2006. Those branches opened toward the end of their valid license window in 2007, and would be included in the January 1st, 2008 branch counts.

I re-estimate the above using a fuzzy RD applying a triangular kernel and present the results in panel A of table 2. These estimates verify that the largest discontinuities begin in 2008, estimated precisely at the 5% and 1% confidence levels. The results for branch licenses, which had been granted

[27]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.

[28]Thanks to Johannes Schmieder for helpful suggestions in displaying the dynamic effect graphically. These figures present estimations that apply a uniform kernel.

but not necessarily yet turned into an operating branch, support the hypothesized waiting-out-the-grace-period strategy described above. The jump in the magnitude of the discontinuities on licenses occurs in 2007, one year earlier than for branches. The discontinuity from licenses generally leads that on branches by one year for 2007-2010 before coming into sync. 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 centers below 50,000 people.

The response observed in operating branches and the corresponding timing of changes in licenses, 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% of the sample mean, reported in brackets 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. The last set of rows in Panel A of table 2 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 behavior is less direct. The reform targeted branch openings but not lending

behavior per se. Other policies such as priority sector lending requirements apply at the bank rather than branch or district level. While the reform generates little direct pressure on lending activity, the theoretical framework in section 3 predicts that anticipation effects could be large. The threat of future competition is expected to induce profit maximizing banks to preemptively expand their credit in under banked areas.

The pattern of expanding discontinuities in total district credit from private sector banks reported in Panel B of table 2 are consistent with this hypothesis. Prior to the policy announcement, districts around the threshold show similar levels of credit on average, and I broadly fail to reject the null hypothesis of continuity at the cutoff. In the post reform period, positive discontinuities expand, reaching substantial magnitudes. The estimates for the number of accounts are estimated with precision at the 5% level in 2007 and 2010, and the 10% level for the other years in 2006-2011. The estimates on the amounts of credit outstanding in millions of rupees are imprecise.

The relative increases in credit in underbanked districts in 2006 and 2007 are consistent with preemptive responses by banks to expected competition. I described the expansion as preemptive because it leads the positive discontinuities found in bank branching, which are first estimated with statistical significance in 2008. Taking the discontinuity in 2006 as a measure of the initial preemptive response, an estimated 6,220 additional credit accounts opened in the under banked districts at the cutoff, which is 48% of the sample mean for districts around the cutoff.[29] The large discontinuity of 10.6 thousand additional accounts estimated in 2007, relative to conjoining years, suggests that preemption motivated by anticipated competition exerts a strong effect. In that year, treated districts received an estimated additional increase of 2.4 licenses but did not yet experience significant branch openings. Future competition and first mover advantage for some districts were therefore focused, while realized entry in subsequent years began dampening the advantage. Slightly

[29]The 2006 credit measures are the first ones following the publication of under banked districts. This early year also leaves banks with the least time to respond through branching. The estimated average treatment effects and sample means around the cutoff refer to the private sector banks.

reduced discontinuities emerge in subsequent years[30] until the jump in 2010 with the surge of branch entry.[31]

The estimates at the total credit levels are still relatively noisy. Lending activity to several sectors (e.g. retail and whole sale trade, construction, mining, etc.) comprise the aggregate measures of credit. If certain sectors depend less on branching access to conduct business, including them may obfuscate the underlying effect. I narrow the focus to credit for direct agriculture and personal loans, which constitute a major portion of private sector bank business and are likely to exert a direct impact on households, which are more likely to rely on branching. The discontinuities estimated from this refined set, reported in the lower half of Panel B in table 2, reflect the findings from the aggregate measures, now with improved precision. Further, positive and sizable discontinuities on credit amounts are now precisely measured at the 5% and 1% levels for years 2005-2007 and 2010-2011.[32]

6.1.2 Public sector banks

Public sector banks show no response to the reform, reported in panel C of table 2. This null result is consistent with the theoretical framework assertion that profit maximizing behavior, with which earlier contributions in the literature document government banks struggle, incentivise credit expansions.

[30]The retraction in 2008 may be explained by the exit of a private bank, Bharat Overseas Bank, through acquisition by a public sector bank in 2007.

[31]The smaller positive, though statistically significant, discontinuity in 2005 arrives before the official announcement of underbanked districts. The online appendix discusses potential hints of the forthcoming reform that may have been available ahead of time. Successful early guessing of vulnerable districts might explain the response, though the statistical significance remains surprising. The composition of banks operating in these districts changed in these early years, resulting from acquisitions of inert private banks by nationalised banks and expansion by aggressive private sector banks. The confluence of these factors may contribute to this finding.

[32]The positive estimated discontinuity in 2004 that is significant at the 10% 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.

6.2 Agriculture

6.2.1 Credit to Agriculture

Over 56% of India's workers engaged in agricultural or related activities in 2001.[33] I next examine the impact of the reform on agricultural credit in less populated areas and agricultural productivity.

The post reform years exhibit positive and statistically significant estimates on credit to agriculture in under banked districts. Panel A of table 3 reports the fuzzy RD estimates 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.[34] The first statistically significant discontinuity in credit to direct agricultural activities emerges in 2005, the year following the Vyas Committee Report. The effect grows in 2006 and diminishes slightly in 2007. A strong effect emerges again in 2009, holding 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 at least 60% of the local mean in all post-reform years. A strong effect on indirect agricultural loans only emerges in 2009.

The uneven pattern in lending merits some discussion. The expansion of credit beginning in 2005 is consistent with the timing of the Vyas Committee, the emphasis on agricultural lending by policy makers and 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 with a variety of potential drivers. 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 subsidized credit program 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.[35]

[33] Population Census of India, 2001.

[34] Percentage change is approximated log-differences of credit amounts from 2001 levels.

[35] The Credit Subvention Scheme operated through NABARD, and exclusively distributed

The growth in both forms of agricultural lending after 2008 in under banked districts might be explained by expanding branch networks. The initial reform, plus a series of policy refinements targeting lower population areas, ramped up new branch openings in later years. A refinement to the branching policy in 2008 created greater incentive for banks to branch into lower populated areas. A new branching policy introduced at the end of 2009 reduced the cost of branch entry in all rural and semi-urban areas, but created additional incentives in under banked districts.

Alternatively, a reform to priority sector lending in 2007 placed greater emphasis on agricultural lending, perhaps leading to more intensive lending through branches. 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 separate government financed debt forgiveness scheme across all commercial banks in June 2008, for small farmers with delinquent direct-to-agriculture loans, may have shuffled new debt free borrowers across public and private banks.[36] Finally, a revision to loan categorization by the RBI in 2008 reduces the accuracy of comparisons by sector around this year. Without finer loan data, disentangling these effects is likely not possible.

6.2.2 Agricultural outcomes

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

through government sector banks, was initiated in 2006-2007.

[36]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 favor them in subsequent years.

of agriculture.[37] 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,[38] 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 2001-2002 to 2009-2010 from data on crop prices and production statistics collected at the district level. The price data for crops is 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 fuzzy RD analysis are shown at the top of Panel B in table 3. The estimates show positive discontinuities of sizable magnitude beginning in 2005, though are estimated imprecisely except for 2009.[39]

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. I present the fuzzy RD results in columns (12) and (13) for the pre-reform and post-reform periods, respectively, 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% level. The

[37]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.

[38]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.

[39]Bias-corrected estimates in the online appendix find statistical significance for 2005, 2009 and 2010.

estimate may be interpreted as every thousand private bank accounts increases the crop yield by an average of 2.3%. 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).[40]

6.3 Industrial Activities

Cities and towns in underbanked districts also experienced significant branch expansion driven by the reform. Manufacturing would be expected to benefit from this improved access to banking. The bottom of Panel A in table 3 presents the fuzzy RD analysis 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.[41] Credit growth from the under banked side emerged beginning in 2008, roughly doubling by 2010 and 2011. Positive discontinuities are estimated with statistically significance in 2009 and 2010. The timing of the effect coincides with the actual opening of branches. The lack of preemptive credit expansion, as observed in total credit, may be attributable to stronger underlying competition in lending to manufacturing, or to the changes in sectoral credit reporting after 2008 noted above. The next section provides evidence from input decisions reported by registered manufacturing firms, including financing.

6.3.1 Evidence from the ASI

In table 4 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

[40]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%, a little less than half of the effect for rainfall found in Jayachandran (2006). Alternatively, leaving cotton out of the index reduces the effect in the post reform period to about 1% and loses statistical significance. In unreported results, using difference-in-differences analysis while limiting the sample of districts to those around the threshold yields positive average effects of the policy on the crop index, and on total revenue from crops, with statistical significance.

[41]Banked districts show little response at the cutoff. See the online appendix.

effect is positive but imprecisely estimated at a value of 17%. The effect on logged working capital, in column (2), is estimated at 0.264 with significance at the 10% level. The effect on the amount of outstanding loans held by the firm in column (3) is estimated to increase 24% with statistical significance at the 10% level. Total investment presented in column (4) increased by 19.7%, with statistical significance at the 10% level. In the last column, the capital-labor ratio is estimated to increase by 3.4 in response to the policy and is also estimated with precision at the 10% significance level. The sample mean of the capital-labor ratio for the under banked states sample was 10.88 post reform, making this a sizable 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 3 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-labor 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.[42] 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.

The bottom of Panel B in table 3 reports the estimated discontinuities. I estimate these using local quadratic regressions as they appear to fit the nightlights data better. I report the bias-corrected measures here, and in the pooled analysis to follow, as these also incorporate information from local quadratic regressions with wider bandwidths, though these closely follow the conventional estimates in the annual regressions. A small negative discontinuity is estimated in 2005, followed by a nearly zero estimate in 2006, with both estimates being small relative to sample mean around the threshold. A positive jump in the discontinuity to 12.9% occurs in 2007 and is estimated significantly at the 1% level. The average change in light for the sample that year increased as well to 11.4%. The discontinuity held constant through 2008, though the average growth in light from 2004 declined, such that the relative magnitude of the discontinuity was greater. Lower levels of light are emitted overall in 2009. The last three years show slightly higher discontinuities than 2007, with statistical significance at conventional levels in 2010 and 2011.

Connecting expanding branch presence to overall economic growth, I perform a fuzzy RD of the change in light on private bank branches. I split the data into a pre-reform period, just 2005 in this case, and a post reform period, 2006-2012. I pool the data within periods and add year fixed effects. I

[42]Comparing changes in light across time for one area to those in another, as opposed to direct cross-sections is important. This helps account for switching satellites, aging of instruments, and annual differences in processing low levels of light.

run the estimation using local linear regressions, because these better fit the bank branching data for a stronger first stage. I still report the bias-corrected estimates, which incorporate local quadratic regressions as discussed. The pre-reform effect reported in column (12) is negative and small, consistent with the earlier findings. The post reform estimate in column (13) is significant at the 1% 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, with each branch increasing nighttime light by 1.15%. Taking the estimated elasticity of nighttime light to GDP from Henderson et al. (2012) of 0.3, this implies that each bank branch raises local GDP by approximately 0.33%. 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%.

6.5 Robustness and Discussion

6.5.1 Robustness to NREGA

I find evidence against the hypothesis that a competing intervention, the Mahatma Gandhi National Rural Employment Guarantee Act (NREGA), drives the observed change in bank branching and subsequent effects. NREGA, coinciding closely in time with the branching reform, constitutes a public works program aimed at relieving poverty by providing 100 days of guaranteed work to individuals in rural areas. Benefits are distributed through bank accounts.[43]

I exploit the phased roll-out of NREGA over 2006-2009, described in Zimmermann (2012) and Klonner and Oldiges (2014). Districts were assigned to the various roll-out phases based on a composite index on district "backwardness." I test whether a discontinuity in district phase assignment can be detected at the cutoff from the branching policy reform. The test fails to reject the null hypothesis of continuity at the cutoff for all three phases. Thus, no

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

difference in NREGA implementation, and therefore benefits, is detected for banked or under banked districts at the cutoff. [44]

6.5.2 Behavior 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 3 may show a steeper negative slope just to the left of the cutoff than further into the set of control districts. While that pattern could 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.[45]

7 Conclusions

This paper analyzes a previously unstudied policy reform in India introduced in 2005, finding new evidence of positive effects from bank branch expansion on economic growth. Bank branching should not be considered a panacea. I attribute the positive effects found here to a set of key factors promoted, intentionally or not, by the interaction of policy incentives and banking environment. However, the results uncover valuable insights. The concentrated response to the reform from private sector banks highlights that banks and their branches act as strategic players responding to incentives. The mobilization of the private banking sector helps explain the positive findings in this

[44] Regression results are available in the online appendix. Additional tests for NREGA are also reported, including a McCrary test for districts excluded from the program.

[45] General equilibrium effects of the reform could be a concern for identifying the unbiased relationship between banking and economic outcomes if, for instance, improved agricultural output in treated districts due to access to finance affects markets in geographically nearby districts otherwise not impacted by the reform. Such impacts that feed back into local credit markets may bias estimates. The potential redistribution of assets following the reform precludes a full discussion of welfare effects. Exploring general equilibrium effects through a fully specified model is left to future work.

work, and bridges 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 underserved 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%. 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 labor ratios. The empirical strategy in my paper identifies these effects independently of growth from the NREGA public work program 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 nighttime light intensity data, estimating that each additional private bank branch led to a 0.33% 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 industrialization, 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 laborers. 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

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

Angelucci, M., D. Karlan, and J. Zinman (2015). Microcredit impacts: Evidence from a randomized microcredit program placement experiment by compartamos banco. *American Economic Journal: Applied Economics 7*(1), 151–82.

Angrist, J. D. and V. Lavy (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 P. Novosad (2012). Factory farming: The impact of agricultural output on local economic activity in India. Technical report, Working Paper.

Banerjee, A. and E. Duflo (2001). The nature of credit constraints: Evidence from an Indian bank. Technical report, MIT Working Paper.

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

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

Banerjee, A. V. and E. Duflo (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., R. Felici, and M. Pagnini (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*, 577–599.

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

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

Calonico, S., M. D. Cattaneo, and R. Titiunik (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 C. Udry (2007). Rural financial markets in developing countries. *Handbook of Agricultural Economics 3*, 2857–2908.

Dehejia, R. and A. Lleras-Muney (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 P. Klemperer (2007). Chapter 31 coordination and lock-in: Competition with switching costs and network effects. Volume 3 of *Handbook of Industrial Organization*, pp. 1967 – 2072. Elsevier.

Field, E., R. Pande, J. Papp, and N. Rigol (2013). Does the classic microfinance model discourage entrepreneurship among the poor? Experimental evidence from India. *American Economic Review 103*(6), 2196–2226.

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

Gelman, A. and G. Imbens (2018). Why high-order polynomials should not be used in regression discontinuity designs. *Journal of Business & Economic Statistics*, 1–10.

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

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

Hsieh, C.-T. and P. J. Klenow (2009). Misallocation and manufacturing TFP in China and India. *The Quarterly Journal of Economics 124*(4), 1403–1448.

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

Imbens, G. W. and T. Lemieux (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 P. E. Strahan (1996). The finance-growth nexus: Evidence from bank branch deregulation. *The Quarterly Journal of Economics*, 639–670.

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

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

Karlan, D. and J. Zinman (2011). Microcredit in theory and practice: Using randomized credit scoring for impact evaluation. *Science 332*(6035), 1278–1284.

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

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

Klonner, S. and C. Oldiges (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., D. Nandy, and M. Puri (2014). Does financing spur small business productivity? Evidence from a natural experiment. *Review of Financial Studies*, hhu087.

La Porta, R., F. Lopez-de Silanes, and A. Shleifer (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 T. Lemieux (2010). Regression discontinuity designs in economics. *Journal of Economic Literature 48*(2), 281–355.

Lee, D. S., E. Moretti, and M. J. Butler (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 M. Rosenzweig (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 L. Zingales (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 A. Weiss (1981). Credit rationing in markets with imperfect information. *The American Economic Review*, 393–410.

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

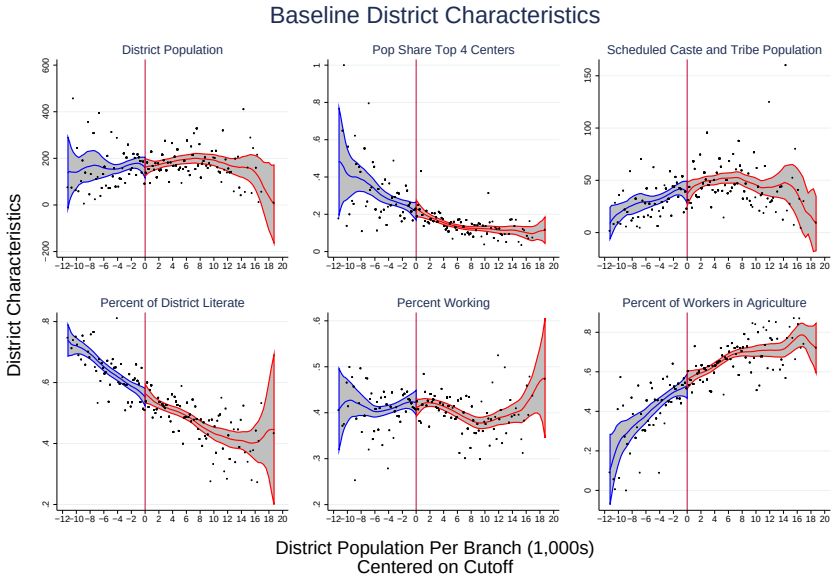
Van der Klaauw, W. (2002). Estimating the effect of financial aid offers on college enrollment: A regression–discontinuity approach. *International Economic Review 43*(4), 1249–1287.

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.

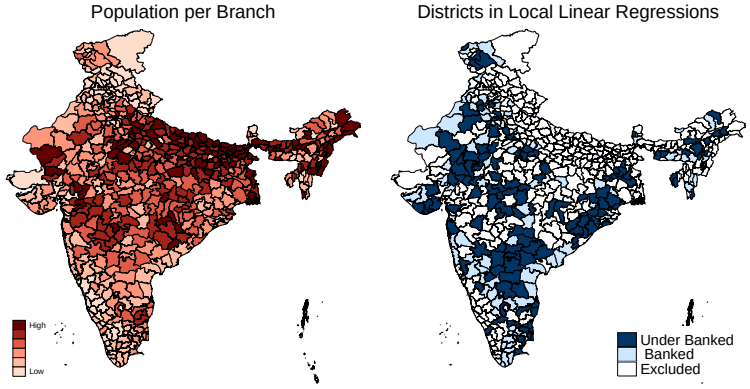
8 Figures and Tables

Figure 1: Continuity Around the Threshold



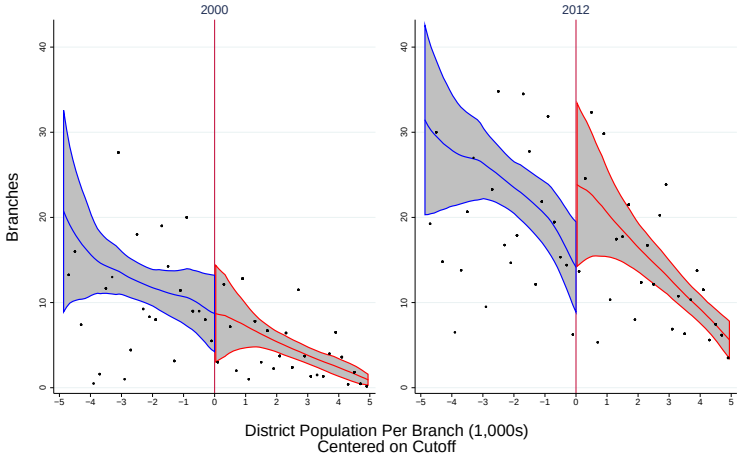
Note: The figure presents baseline district characteristics taken from the 2001 Population Census of India, with dots reporting local averages for districts falling within non-overlapping 200 persons per branch bins. The horizontal axis is the forcing variable of district population per branch centered 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 3.5 thousand persons per branch is shown at each x-value, allowing for different slopes on either side of the cutoff, with 5% confidence intervals.

Figure 2: Heat Maps of District Population Per Branch and Treatment Status



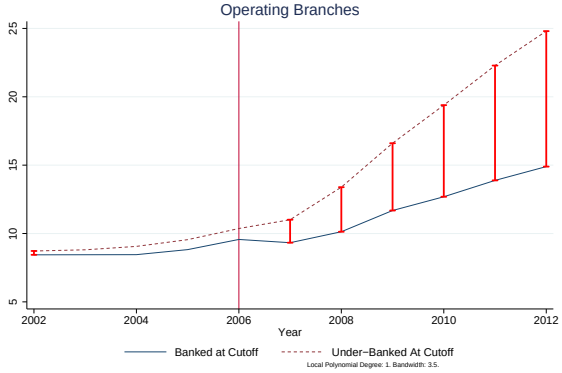
Note: Heat map of district population per branch is on the left. District under banked status, excluding districts outside local linear regressions bandwidth, is on the right.

Figure 3: Visual RD: Operating Private Bank Branches



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 3.5 thousand persons per district bandwidth and triangular kernel, at each x-value and the 5% 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 4: Discontinuity from Reduced Form by Year: Private Bank Branches



Note: Estimated using local linear regressions with controls for district population and its square, the percent of workers in agriculture and the pre-randomization 2001 value of the dependent variable. Bandwidths are set 3.5 thousand 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% level. The thin vertical line at 2006 represents the first estimation made following the reform.

Table 1: Summary Statistics

	Banked, Pre-reform			Banked, Post-reform			Under Banked, Pre-reform			Under Banked, Post-reform		
	N	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD
Panel A: Banking												
Branches:												
SBI	610	28.62	23.10	732	33.24	27.97	900	21.35	16.38	1,080	24.60	19.69
Nationalised	610	69.81	62.76	732	80.63	73.98	900	45.44	44.86	1,080	51.54	50.43
RRB	610	21.52	21.68	732	23.00	22.17	900	28.22	22.15	1,080	29.21	22.95
Foreign	610	0.12	0.62	732	0.24	1.00	900	0.02	0.20	1,080	0.11	0.46
Old Private	610	11.81	16.58	732	11.63	15.30	900	4.20	9.30	1,080	4.59	10.00
New Private	610	2.43	4.60	732	7.25	10.69	900	0.79	2.37	1,080	4.05	6.15
Public Banks	610	120.15	87.59	732	137.93	105.53	900	95.06	66.49	1,080	106.01	76.43
Private Banks	610	14.23	18.58	732	18.88	20.76	900	4.99	10.38	1,080	8.64	13.96
Credit Amount:												
SBI	610	5,293.64	5,980.07	732	11,037.75	12,248.84	900	3,285.65	5,986.45	1,080	6,507.47	8,548.99
Nationalised	610	10,236.99	13,154.39	732	22,228.23	33,180.44	900	4,602.58	5,692.05	1,080	9,362.26	12,494.34
RRB	610	870.75	1,198.64	732	1,738.28	2,270.79	900	950.14	1,134.78	1,080	1,869.28	2,256.91
Foreign	610	201.34	727.79	732	487.36	1,620.56	900	50.17	293.39	1,080	191.79	1,414.19
Private	610	3,813.91	7,071.33	732	7,637.43	12,055.83	900	1,354.92	3,542.47	1,080	2,437.96	5,464.27
Credit Accounts:												
SBI	610	30,945.37	31,517.42	732	47,639.10	50,181.88	900	24,107.01	24,218.30	1,080	38,046.44	39,105.76
Nationalised	610	60,582.56	60,584.96	732	89,278.02	97,041.33	900	37,964.00	38,526.22	1,080	55,938.74	58,976.20
RRB	610	22,255.54	33,920.33	732	30,088.21	47,295.12	900	28,251.07	34,646.61	1,080	36,354.23	48,093.88
Foreign	610	134.43	772.63	732	319.86	1,656.41	900	51.02	564.60	1,080	119.10	874.72
Private	610	9,792.66	14,751.41	732	25,507.24	3,5027.74	900	3,214.42	7,356.89	1,080	9889.30	22,363.60
Deposit Amount:												
SBI	607	9,599.80	10,660.29	732	16,412.71	20,661.42	892	6,104.53	6,197.59	1,078	10,180.87	10,886.09
Nationalised	607	20,027.74	26,126.93	732	33,469.46	51,159.49	892	9,745.18	12,975.67	1,078	15,306.32	20,677.41
RRB	607	1,340.93	1,519.90	732	2,212.51	2,520.01	892	1,807.67	1,792.85	1,078	2,828.68	2,818.40
Foreign	607	181.20	1,207.17	732	611.75	4,849.06	892	20.19	243.41	1,078	65.09	603.55
Private	607	4,695.24	8,722.10	732	8,973.14	17,799.64	892	1,371.38	2,938.95	1,078	2,798.10	5,257.67
Deposit Accounts:												
SBI	607	203.44	178.68	732	298.25	276.02	892	147.73	130.51	1,078	232.13	228.96
Nationalised	607	502.83	502.30	732	683.75	696.66	892	294.64	342.96	1,078	410.49	464.15
RRB	607	76.55	101.22	732	118.76	157.80	892	100.52	109.82	1,078	157.42	174.79
Foreign	607	0.98	6.61	732	2.27	14.99	892	0.19	2.34	1,078	0.40	2.83
Private	607	91.00	124.47	732	136.98	184.15	892	30.16	63.39	1,078	50.57	93.78
Panel B: Annual Survey of Industries												
Log Total Employees	42,702	3.79	1.40	40,252	3.95	1.44	21,133	3.57	1.35	17,976	3.72	1.40
Log Number of units	42,824	0.04	0.19	40,575	0.04	0.20	21,216	0.02	0.15	18,123	0.03	0.16
Plant Age	42,248	16.00	13.99	39,268	15.20	13.88	20,864	14.97	14.20	17,562	14.66	14.33
Log Capital (No Land or Inventory)	42,339	14.91	2.88	39,707	15.15	3.39	21,030	14.58	2.95	17,886	15.00	3.14
Log Net Assets	42,352	15.68	2.88	39,772	15.76	3.29	21,040	15.35	2.93	17,902	15.60	3.02
Log Working Capital	35,823	15.31	3.02	34,057	15.26	3.69	18,262	15.02	3.11	15,818	15.29	3.15
Log Loans	34,828	14.87	4.04	32,543	14.96	4.20	16,258	14.87	4.08	13,795	15.06	4.04
Log Total Investment	39,950	14.69	3.20	37,858	14.94	3.83	20,517	14.25	3.30	17,468	14.65	3.62
Capital Labor Ratio	42,221	6.64	47.52	39,543	11.12	237.38	20,971	8.13	38.90	17,800	10.88	105.47
Log Capital Labor Ratio	42,202	0.77	1.54	39,535	0.88	1.52	20,958	0.89	1.66	17,798	1.00	1.65

Note: Panel A. Sources include the RBI Master Office File, the BSR 1 and BSR 2 for the years 2001-2011. The sample includes districts falling within 5,000 persons per branch of the cutoff. Each year includes 122 banked districts and 180 under banked districts from a total of 572 considered districts. Amounts are reported in millions of Rupees; accounts are reported in thousands. Panel B. Annual Survey of Industries, Unit level data 1999-2010. The sample includes plants reporting being open and providing a valid urban or rural status. The capital-labor-ratio is constructed as the average of the opening and closing Net Assets divided by the total wage bill inclusive of benefits. States were selected according to the concentration of their population around the cutoff. "Banked states" include Haryana, Uttarakhand, Punjab, Mizoram, Daman and Dimu, Karnataka, Puducherry. "Under banked states" include Rajasthan, Tripura, Jharkhand, Orissa, Dadra and Nagar Haveli.

Table 2: Fuzzy RD of Banking Outcomes: First Stage Estimates Underbanked Assignment

	(1) 2002	(2) 2003	(3) 2004	(4) 2005	(5) 2006	(6) 2007	(7) 2008	(8) 2009	(9) 2010	(10) 2011	(11) 2012
Panel A: Bank branches											
Licences	-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)
	[8.714]	[8.917]	[9.241]	[9.847]	[10.62]	[11.92]	[13.83]	[15.31]	[17.13]	[18.47]	[19.99]
Operating Branches	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)
	[8.636]	[8.801]	[9.125]	[9.597]	[10.34]	[10.87]	[12.25]	[14.42]	[16.19]	[17.91]	[20.00]
Operating Banks	-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)
	[2.788]	[2.963]	[3.171]	[3.532]	[3.991]	[4.269]	[4.787]	[5.421]	[5.875]	[5.856]	[6.250]
Panel B: Private Sector Credit											
Accounts: Aggregate	-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)	–
	[5.067]	[5.484]	[6.470]	[8.800]	[12.83]	[13.77]	[16.14]	[17.78]	[22.82]	[25.80]	
Amounts: Aggregate	-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)	–
	[2,641]	[3,223]	[2,943]	[3,466]	[3,922]	[4,920]	[5,278]	[5,362]	[4,932]	[5,990]	
Accounts: DA & PL	-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)	–
	[3.311]	[3.460]	[4.640]	[6.357]	[9.937]	[9.638]	[10.82]	[13.69]	[15.78]	[17.77]	
Amounts: DA & PL	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)	–
	[236.9]	[309.6]	[777.2]	[1,209]	[1,603]	[1,978]	[1,989]	[2,121]	[1,856]	[2,269]	
Panel C: Public Sector Credit											
Accounts: Aggregate	-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)	–
	[99.51]	[102.5]	[104.7]	[120.6]	[132.4]	[141.5]	[150.5]	[154.2]	[167.2]	[177.1]	
Amounts: Aggregate	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)	–
	[10,586]	[11,953]	[13,479]	[17,693]	[21,386]	[23,326]	[27,222]	[29,581]	[31,372]	[34,125]	

Note: Standard errors in parenthesis, sample mean around the cutoff 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 percent of workers in agriculture, a control for monsoon rainfall and the pre-randomization 2001 value of the dependent variable. Bandwidths are set 3.5 thousand persons per branch and estimated using a triangular kernel. Under banked status is instrumented for with predicted under banked assignment. **Panel A.** Licenses are considered in operation if they are granted for a branch currently operating or pending opening. **Panel B.** Accounts reported in thousands. Amounts reported in millions of rupees. DA and PL refers to summed direct to agriculture and personal loans. **Panel C.** Public sector banks include State Bank of India and Associated Banks, Nationalised Banks, IDBI and Regional Rural Banks.

Table 3: Fuzzy RD of Sector and Real Economy Effects: First Stage Varies

	(1) 2002	(2) 2003	(3) 2004	(4) 2005	(5) 2006	(6) 2007	(7) 2008	(8) 2009	(9) 2010	(10) 2011	(11) 2012	(12) Pre	(13) Post
Panel A: Credit to Sectors													
Percentage change in credit amounts to Agriculture													
Direct	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)	–	–	–
	[-0.0700]	[0.0481]	[0.164]	[0.433]	[0.550]	[0.964]	[1.488]	[1.419]	[1.953]	[2.376]			
Indirect	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)	–	–	–
	[0.317]	[0.199]	[0.188]	[0.237]	[0.257]	[0.453]	[0.955]	[1.039]	[1.313]	[1.133]			
Percentage change in credit amounts to Manufacturing and Processing													
	-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)	–	–	–
	[0.934]	[1.098]	[0.678]	[0.763]	[0.553]	[0.694]	[1.119]	[1.231]	[1.287]	[1.410]			
Panel B: Real Sector													
Crop Yield Index	0.105	-0.0327	-0.0324	0.250	0.160	0.158	0.0846	0.353*	0.251			-0.0412	0.0230*
	(0.162)	(0.205)	(0.194)	(0.170)	(0.194)	(0.266)	(0.225)	(0.204)	(0.191)	–	–	(0.168)	(0.0138)
	[0.121]	[-0.0216]	[0.106]	[0.117]	[0.119]	[0.0945]	[0.107]	[0.0857]	[0.0860]			[0.0685]	[0.102]
Percentage change				-0.0430*	0.00773	0.129***	0.128***	0.0567	0.151*	0.141**	0.148	-0.0455	0.0115***
in Mean District	–	–	–	(0.0227)	(0.0325)	(0.0415)	(0.0460)	(0.0778)	(0.0773)	(0.0652)	(0.105)	(0.274)	(0.00373)
Light				[-0.139]	[-0.0805]	[0.114]	[0.0722]	[0.0266]	[0.355]	[0.219]	[0.297]	[-0.139]	[0.143]

Note: Standard errors in parenthesis, sample mean around the cutoff in brackets. 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 percent of workers in agriculture and a control for monsoon rainfall. Bandwidths are set 3.5 thousand persons per branch and estimated using a triangular kernel. **Panel A.** Percentage change is approximated using difference in logs relative the value reported in 2001. Under banked status is instrumented for with predicted under banked assignment. Credit amounts to agriculture are for use in semi-urban and rural areas. Amounts for manufacturing are for all population groups. **Panel B.** Crop yield index uses weighted averages of the crops rice, wheat, jowar, groundnut and cotton. Weighted by crop revenue share. Year fixed effects are added and the pre-randomization value is dropped. The number of private sector bank accounts is instrumented for with predicted under banked assignment. Pre-reform years include 2002-2004 and post-reform is 2005-2010. Nighttime lights: Annual estimates use local quadratic regressions and the usual controls. The fuzzy regression discontinuity is estimated using local linear regressions. The number of operating private bank branches is instrumented with predicted under banked assignment. Pre-reform year is 2005 using 2004 as the base year for the approximate percentage change. Post-reform years are 2006-2012.

Table 4: Diff n Diff: States Selected around Under Banked Threshold, 1999-2010

VARIABLES	(1) Log Net Assets	(2) Log Working Capital	(3) Log Loans	(4) Log Total Investment	(5) Capital 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

Note: Robust standard errors in parenthesis. *** p<0.01, ** p<0.05, * p<0.1 Standard Errors Clustered at State level. 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.

9 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 harmonization 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:

9.1 Constructing the forcing variable

In constructing the forcing variable and national average I follow the APPBO procedure [46] 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 September 7th, 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.[47] Out of 572 districts only 5 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 form from two or more source districts,

[46] 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 specialized bank business which may depart from the actual algorithm used by the RBI. The national average to which the value is compared is the total population of India divided by the number of bank offices.

[47] 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 September 7th, 2005 as the policy date was 14,915 persons per branch in India.

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 January 1st, 2006 yields similar results.

9.2 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 sugar production in order to match the two data sets, and whether the reported prices for sugar capture actual prices faced by farmers after accounting for delay of payments bargaining. Crop yields are normalized 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 data sets, 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 2002-2010.