Bank Presence and Health

By Kim Fe Cramer*

This paper examines whether more bank presence in underserved areas can improve households' health. I utilize a policy of the Reserve Bank of India from 2005, applying a regression discontinuity design. Five years after the policy introduction, treatment districts have 27 more branches than control districts. This improved financial access in treatment districts positively impacts health. Six years after the policy, households are 19 percentage points less likely to suffer from a non-chronic illness in a given month. Chronic diseases remain unaffected. I provide suggestive evidence that an employment effect, household savings accounts, and hospital credit contribute to the impact. In contrast, personal bank loans play no role for the average household.

Keywords: Financial Development, Banks, Health
JEL Codes: G21, O16, I15

*LSE Finance Department, Houghton St, London WC2A 2AE, United Kingdom (email: k.f.cramer@lse.ac.uk, www.kimfecramer.com). Supported by the Chazen Institute for Global Business and the Bernstein Center for Leadership and Ethics at Columbia University. I thank Giorgia Barboni, Emily Breza, Miriam Bruhn, Robin Burgess, Xavier Giroud, Sean Higgins, Jonas Hjort, Dean Karlan, Nicola Limodio, Michaela Pagel, Rohini Pande, Suresh Sundaresan, Nishant Vats, Eric Verhoogen, Jack Willis, Daniel Wolfenzon, and participants at the American Economics Association Meeting and other conferences and seminars for their valuable comments. Additionally, I am grateful to International House New York for its inspiring environment.

What are the general equilibrium effects of banks on households? Previous research has focused on labor markets. Banks extend credit to firms, thereby fostering economic activity and employment (Bruhn and Love, 2014; Fonseca and Matray, 2022). In this paper, I demonstrate that beyond that, bank presence can contribute towards tackling hard-to-crack development challenges, studying the third UN Sustainable Development Goal of improving households' health. In addition to stimulating employment that allows households to invest more in health, banks may improve health through three distinct activities. First, banks might offer savings accounts to households. Second, they may provide personal bank loans to households. Both savings accounts and bank loans could allow households to invest more in health when necessary. Third, banks could extend credit to healthcare providers, thereby stimulating healthcare supply, a crucial determinant of health status. Despite these strong motivations, we lack causal evidence on the impact of bank presence on households' health.

To obtain exogenous variation in bank presence, I use a Reserve Bank of India (RBI) policy from 2005. The policy incentivizes banks to set up new branches in underbanked districts. These districts have a population-to-branch ratio that exceeds the national average. In a regression discontinuity design, I compare districts with a ratio just above the national average (treated) and those just below (control). I use RBI data to test whether more branches are set up in treatment districts. To measure health impacts, I utilize two nationally representative household surveys, the Indian Human Development Survey (IHDS) six years after the policy introduction and the Demographics and Health Survey (DHS) ten years after. To provide supplementary evidence on banking activities, I utilize information on households' financial access from the IHDS and information on firms and healthcare providers from the Economic Census.

Initially, I demonstrate that the policy increased bank presence in treatment districts. Smooth before the policy introduction, treatment districts have significantly more branches two years later. Matching the dynamics of the policy, these discontinuities continue to grow. Five years post-policy, treatment districts have 19 percent or 27 more branches, compared to 142 branches in control districts. This effect is economically meaningful. Utilizing administrative data on deposit accounts and total credit, I find that treatment districts have 161,977 more deposit accounts and 148 million USD more in credit. Moreover, private banks were the driving force behind the expansion. While 17% of branches are private in the control, 53% of the new branches are private. Entry of private banks could have further boosted competition in the banking sector. Overall, this suggests that the policy introduced exogenous and economically meaningful bank entry.

Following the branch expansion, I detect positive impacts on households' health. Six years after the policy's implementation, households' probability of suffering from a non-chronic illness such as fever or diarrhea in a given month is 19 percentage points lower, compared to a control mean of 52 percent. This effect size is in the middle of the range of other successful health interventions in developing countries (see Waddington et al. (2009) for a meta-analysis). The reduction in non-chronic diseases positively affects labor supply and school attendance. I do not find effects for chronic diseases such as diabetes. The second survey conducted a decade after the policy introduction allows me to replicate my findings on non-chronic illnesses. Thus, evidence from two different surveys demonstrates that bank presence can play a vital role in improving health.

I provide extensive evidence to reject potential threats to causal identification. First, I establish that local governments do not manipulate their treatment status. By construction, manipulation of the population-to-branch ratio is unlikely. The numerator relies on historical data from the 2001 Population Census. In the denominator, the total number of branches is the sum of individual decisions of all banks in a district. Additionally, banks directly report their number of branches to the RBI. Indeed, I find no evidence that more districts are located just above than just below the cutoff. Nor is there any evidence that districts just above and below the cutoff significantly differ before the policy. To demonstrate this, I utilize data from pre-policy rounds of the IHDS, the Economic Census, and the Population Census, as well as night-light data. There is also no threat to identification due to migration, which is negligible. Finally, no policies use an identical cutoff or are significantly more likely to be implemented in treatment districts. Results are robust under varying bandwidths and polynomials, and there is little evidence of discontinuities at placebo cutoffs. In summary, this evidence strengthens the confidence in the causal interpretation of my findings.

Supplementary to my main results, I provide suggestive evidence of specific banking activities. I find empirical evidence for three out of four banking activities. First, in alignment with previous studies, banks interact with businesses, generating positive employment effects. Second, banks provide savings accounts to households. Third, banks offer credit to healthcare providers, which, in equilibrium, expand healthcare supply. In contrast to these three banking activities, I do not find evidence that households gain access to personal bank loans. The coefficient is small and insignificant. This contradicts the narrative that households utilize formal medical debt. Thus, bank presence likely affects health through established activities (firm credit, household savings accounts) and understudied aspects (credit to healthcare providers), but not through personal bank loans. Readers curious about whether one specific banking activity in isolation would have been sufficient to improve health can turn towards randomized controlled trials. This paper instead takes a comprehensive view, relaxing the financial constraints of multiple actors in the economy simultaneously. This acknowledges the complexities of healthcare markets, where banks can stimulate both demand- and supply-side. Furthermore, it informs branch policies implemented globally, including in countries like Brazil and China.

How did these banking activities translate into health improvements? To investigate this further, I study households' health-related investments. Households might improve their health by spending more on low-fixed-cost items such as food or hygiene, high-fixed-cost items such as toilet facilities or fridges, or increasing healthcare demand. Consistent with a gradual increase in the availability of resources through employment and savings accounts, I find empirical evidence that households spend more on food and hygiene. In alignment with a lack of credit take-up, I do not find evidence that households invest in high-fixed-cost items. Finally, I provide suggestive evidence that households increase their healthcare demand. To summarize, banking activities likely affected health through increasing households' spending on low-fixed-cost items and healthcare demand, but not through households' investments in high-fixed-cost items.

The contribution of this study is to examine the impact of bank presence on health. It primarily informs the literature on the general equilibrium effects of bank presence on households. This literature includes studies in developing countries (Burgess and Pande, 2005; Bruhn and Love, 2014; Barboni et al., 2021; Fonseca and Matray, 2022) and developed countries (Brown et al., 2019; Célerier and Matray, 2019; Stein and Yannelis, 2020). While prior work has established that bank presence can stimulate employment and household income, it has not explored the relationship to health. One might raise the question of whether we can simply extrapolate that health must improve. We might be inclined to do so if there was a strong relationship between income and health in developing countries. Empirically, substantial cash transfers to households do not show positive health effects (Haushofer and Shapiro, 2013; Egger et al., 2019). Theoretically, two key factors could drive a wedge between income and health. First, households might not spend more on health due to non-monetary transaction costs, lack of information, or behavioral biases (Dupas and Miguel, 2017). Second, even if households spend more on health, the healthcare supply might not sufficiently expand to improve health if there are high fixed costs to investments and credit constraints. Therefore, we need to evaluate the impact on health independently.

This paper closely connects to a second literature that studies the impact of other forms of financial access in developing countries (Kanz, 2016; Agarwal et al., 2017; Giné and Kanz, 2018; Limodio, 2022; Higgins, 2020; Bachas et al., 2021;

Breza and Kinnan, 2021; Doornik et al., 2021; Garber et al., 2021; Aydin, 2022; De Roux and Limodio, 2023; Fonseca and Van Doornik, 2022; Ghosh and Vats, 2022; Fiorin et al., 2023). Whether these other forms of financial access can affect health has been primarily explored through randomized controlled trials that offer financial products to households. These studies frequently find null results for savings accounts (Dupas and Robinson, 2013; Prina, 2015; Dupas et al., 2018), bank credit (Karlan and Zinman, 2010), and microcredit (Beaman et al., 2014; Angelucci et al., 2015; Attanasio et al., 2015; Augsburg et al., 2015; Banerjee et al., 2015b; Crépon et al., 2015; Tarozzi et al., 2015; Banerjee et al., 2019). There are two important distinctions between my work and these randomized controlled trials. First, examining bank presence allows me to capture the effects of finance that are mediated through the labor and healthcare market. Second, a large treated sample and long-term effects up to ten years after the policy introduction allow me to capture general equilibrium effects.

Finally, this paper speaks to a growing literature that connects finance and health in developed countries. This literature explores, for instance, the relationship between households' financial decisions and health, such as the mental health effects of credit (Andersen et al., 2022). It also investigates the relationship between hospitals' finances and health, for example, the effect of a financial crisis (Adelino et al., 2022), cash flow shocks (Adelino et al., 2015), credit shocks (Aghamolla et al., 2023), bankruptcies (Antill et al., 2023), and private equity ownership (Gupta et al., 2023; Liu, 2022) on quality of hospital services. These studies demonstrate an increasing interest in the understudied relationship between finance and health.

My findings carry implications for both policy and future research. Policymakers worldwide have been implementing branch-opening initiatives in underserved regions. While we recognize their positive impact on businesses and labor markets, we now learn they also have a role in improving households' well-being beyond their financial situation. This paper also encourages further exploration into the effects of bank presence on various dimensions of well-being, including education. Gaining insights into these inquiries can substantially advance our understanding of the impact of bank presence and the potential for policymakers to enhance their citizens' well-being.

I. Policy

I use a policy the Reserve Bank of India introduced in 2005 to incentivize banks to open new branches in underserved locations. The policy is still in effect and states that banks can increase their chance of obtaining licenses for branches in favored locations by strengthening their branch presence in underbanked districts. Districts are defined as underbanked if their population-to-branch ratio exceeds the national average. In 2006, the RBI published a list of underbanked districts to assist banks in identifying them. District-level ratios are not included in this document, so I reconstruct them as described in Section II. The list of underbanked districts has remained constant since its release; the RBI has not adjusted the list to account for changes in the ratio. Thus, for this study, I employ the cross-sectional variation in the district-level population-to-branch ratio in 2006. In 2010, the RBI adapted its policy to allow branch openings without licenses in eight of the 35 states or union territories that were particularly disadvantaged. I do not exploit this variation for identification, but it appears in the dynamic patterns of banks' responses to the policy. Figure 1 depicts all 593 districts as of the 2001 Census. Marked in dark blue are the 375 districts defined as underbanked according to the reconstructed district-level ratio in 2006.

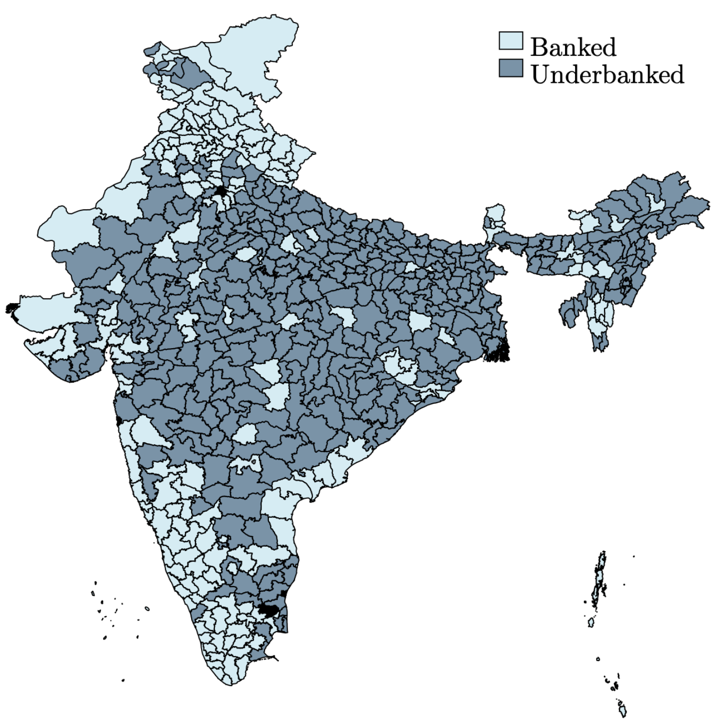


Figure 1. Banked and Underbanked Districts. District borders refer to the 2001 Census.

$$(1) \qquad \underbrace{\frac{\text{Population}_{District}}{\text{\# Bank Branches}_{District}}}_{\text{Underbanked/Treated}} > \frac{\text{Population}_{National}}{\text{\# Bank Branches}_{National}}$$

To my knowledge, this is the first paper that combines the 2005 RBI policy with household-level data. The policy has been utilized in one previous study by Young (2017), who focuses on aggregate outcomes to study the impact on economic activity. Similar to the 2005 policy, another branch licensing policy was in place between 1977 and 1990. Burgess and Pande (2005) use the 1977-1990 policy in their seminal paper on the impact of bank presence on poverty, employing an instrumental variable strategy. The authors focus on state-level measures of poverty. From 1990 through 2005, no comparable branch licensing policy was in place.

II. Data

Initially, I reconstruct the policy's population-to-branch ratio. To measure the population of each district, I rely on the 2001 Population Census (ORGCC, 2008). To measure the district-level number of branches in the denominator, I use an annual publication of the RBI, the Bank Branch Statistics (RBI, 2018a). I focus on data from the first quarter of 2006 since the final list of underbanked districts was issued in July of that year. To measure how banks reacted to the policy, I use a second district-level branch data set: the Master Office File (RBI, 2018b). This file is dynamically updated over time to reflect changes in district borders, which means that when I trace back data to the 2001 Census borders that are used for the policy, I lose accuracy. Thus, I do not use the Master Office File to construct the ratio. The main advantage of this data is that it allows me to study the reaction of different bank types separately. One specific bank type, regional rural banks, is excluded from the policy; correspondingly, I also exclude this bank type from my outcomes of interest. Instead, I utilize regional rural banks to conduct placebo tests. For the years 1997 to 2016, I obtain two variables for all other bank types: the number of branch licenses and the number of branches. Using this data from 1997 to 2004, I test for pre-policy smoothness in bank licenses and branches around the policy cutoff. Data from 2005 to 2016 allows me to examine the respective discontinuities after the policy. In 2016, the final household-level survey was conducted. I supplement this with data on deposit accounts and total credit (RBI, 2018c). General summary statistics from the Master Office File are provided in Table A1.

To examine the effect of bank presence on health, I use two nationally representative household surveys. The first is the Indian Human Development Survey

(IHDS). This panel survey was conducted in 2004/2005 (IHDS I), shortly before the policy, and again six years after the policy in 2011/2012 (IHDS II) (see Figure 2) (Desai and Vanneman, 2018a,b). The pre-policy round allows me to test for the smoothness of household characteristics around the cutoff. The post-policy round provides the primary outcome variables. The survey not only contains health information but also provides a picture of the households' economic situation. With this data, I can test, for instance, how many days of work or school households missed due to illness or whether they hold financial instruments. The first survey round was conducted in 64 percent of districts and the second in 65 percent. Figure A1 depicts districts covered in the second survey round, distinguishing between the 218 underbanked and 166 banked districts. Both survey rounds cover all states and union territories of India except Lakshadweep, and Andaman and Nicobar Islands. The survey was not more likely to be conducted in underbanked districts (see Discussion A1). In the first survey round, 41,554 households were interviewed. In the second round, 83% of the original households plus replacement households were interviewed. This attrition does not threaten identification, as I rely on comparing households in treatment and control districts in the second survey round. General summary statistics of the IHDS are described in Table A2. In this table, I also provide evidence of the external validity of my design; households in districts with a ratio in a range of $\pm 3,000$ of the policy cutoff are very similar to all households in the sample.

I complement the IHDS with a second nationally representative household-level survey, the Demographics and Health Program (DHS), conducted in 2015 and 2016, ten years after the policy (see Figure 2) (IIPS and ICF, 2017). In contrast to the IHDS, the DHS primarily focuses on health. The survey was conducted in all districts and interviewed 601,509 households. The previous round of this survey, conducted in 2005 and 2006, does not contain district-level identifiers. Consequently, I do not include that survey round in my analysis. General summary statistics for the DHS are provided in Table A3.

Complementing my main results, I provide suggestive evidence of specific bank activities. I observe households' financial access in the IHDS (2011/2012). To understand banks' relationship with businesses and healthcare providers, I additionally examine the Economic Census. This data allows me to test for an employment effect and to investigate activity in the healthcare sector. The Economic Census covers all informal and formal establishments in India, except those engaged in activities of farming, plantation, public administration, and defense. I focus on two census rounds; the first was conducted in 2005 and the second in 2013 (see Figure 2) (CSO and MOSPI, 2018a,b). The first Economic Census round allows me to test for smoothness around the cutoff in the respective variables pre-policy. The second round provides outcome variables. Summary statistics are provided in Table A4. To gain a better understanding of the healthcare sector, I investigate summary statistics from the Prowess database, which provides financial statements for companies of all sizes, including those conducting health services (CMIE, 2020). The Prowess sample includes all companies traded on the National Stock Exchange and the Bombay Stock Exchange, as well as a selected sample of unlisted public limited companies and private limited companies. While providing more detailed financial information than the Economic Census, due to the limited number of districts represented in Prowess, I concentrate on the Economic Census for the regression analysis.

To provide further evidence on pre-policy smoothness along other dimensions, including economic activity and population characteristics, I utilize the Socioeconomic High-Resolution Rural-Urban Geographic Data Platform (SHRUG) (Henderson et al., 2011; Asher and Novosad, 2019; Asher et al., 2021). This platform combines multiple data sources on the village or town level. Economic activity is proxied by night-light data, economic employment, and road connections. Population characteristics include total population and literate population.

A final point to note is that India's district borders are very dynamic. While the 2001 Census contains 593 districts, the 2011 Census contains 640 districts (ORGCC, 2014). The RBI policy refers to the 2001 district borders. In contrast, most data sources I use are adjusted for any changes in district borders at the respective time of publication. To analyze treatment effects for districts as defined by the policy, I trace all data back to the 2001 Census borders. The main source for this is the 2011 Census.

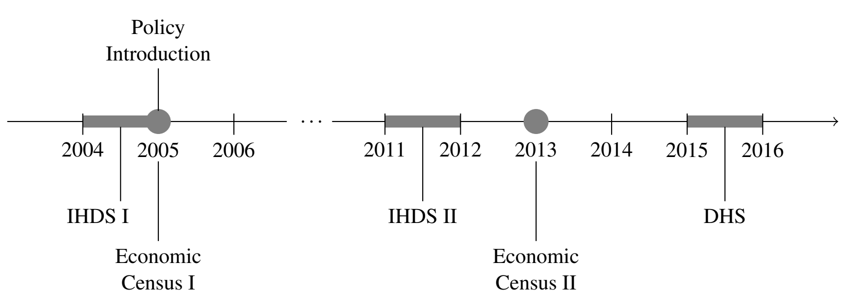


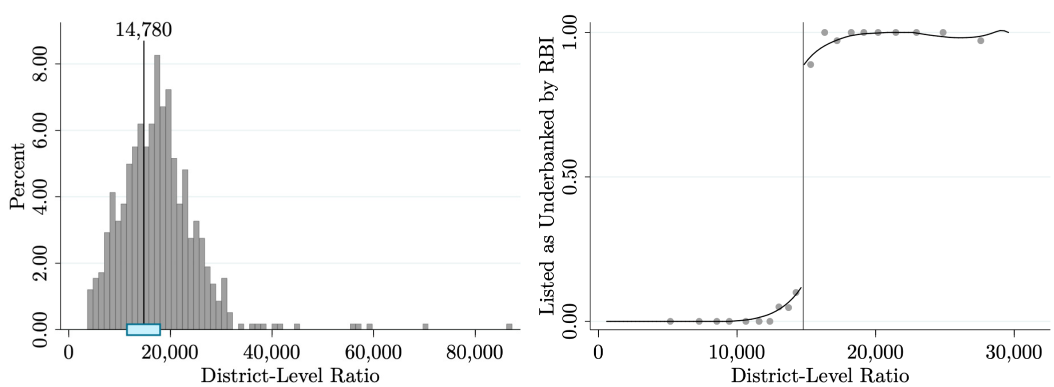
Figure 2. Timeline. The following graphic depicts a timeline of this study, with the three main data sets used (IHDS, DHS, and Economic Census).

III. Identification Strategy

A. Regression Discontinuity Design

The design of the RBI policy allows for a regression discontinuity analysis. The district-level population-to-branch ratio is the running variable, and the national average ratio is the cutoff. Districts with a ratio above the national average are defined as underbanked or treated, while those below the national average are defined as banked or control. Figure 3(a) depicts the histogram of the district-level ratio. The vertical line indicates the national average of the ratio: 14,780. The regression discontinuity analysis concentrates on observations within an optimal bandwidth. While this optimal bandwidth depends on the specific outcome variable (Cattaneo and Vazquez-Bare, 2017), districts included are mostly within a range of $\pm 3,000$ relative to the cutoff. This range is indicated by the blue bar on the x-axis in Figure 3(a). Figure A2 provides a map of districts in this range. As discussed below, for the identification assumption to hold, there should be no perfect manipulation around the cutoff, one implication of which is that there are approximately the same number of districts just above and just below the cutoff. At first glance, the histogram does not appear to show more districts just above the cutoff than just below. I test this formally using the McCrary (2008) density test.

While I do not perfectly predict which districts are listed as underbanked by the RBI, there are only a few districts, 10 out of 593, that have a different status than predicted. There are two potential reasons why I do not perfectly predict which districts are listed as underbanked. First, despite conversations with the RBI, I do not



(a) Histogram (b) First stage

Figure 3. Histogram and First Stage. The vertical line in both graphs indicates the national average of the population-to-branch ratio (14,780).

know which exact data sources they used to construct the ratio. Second, the RBI might have used discretion, deciding to include a district in the list despite having a ratio below the cutoff or vice versa. Both reasons do not threaten identification but give rise to the fuzzy RDD. Figure 3(b) shows that when a district's ratio crosses the national average, there is a large jump in the probability that it is listed as underbanked. Consequently, I implement a fuzzy regression discontinuity design with a strong first stage. I use the following specification for household-level regressions. Regressions on more aggregated levels, such as the district level, exactly mirror the household-level regressions but with higher-level indices.

$$\begin{aligned} \text{Underbanked}_{d,s} = \alpha_0 + \alpha_1 \text{Above}_{d,s} + \alpha_2 \text{DistRatio}_{d,s} \\ + \alpha_3 \text{DistRatio}_{d,s} \text{Above}_{d,s} + \lambda X_{d,s} + \mu_s + \upsilon_{d,s} \end{aligned} \tag{2}$$

$$\begin{aligned} \text{y}_{h,d,s} = \beta_0 + \beta_1 \text{Underbanked}_{d,s} + \beta_2 \text{DistRatio}_{d,s} \\ + \beta_3 \text{DistRatio}_{d,s} \text{Above}_{d,s} + \gamma X_{d,s} + \eta_s + \varepsilon_{h,d,s} \end{aligned} \tag{3}$$

Here h denotes household, d denotes district, and s denotes state. $\text{Underbanked}_{d,s}$ is an indicator equal to one if the district is listed as underbanked. $\text{DistRatio}_{d,s}$ is the district-level ratio. $\text{Above}_{d,s}$ is an indicator equal to one if the district-level ratio is larger than its national average. I control for the ratio's components in $\text{X}_{d,s}$ and include state-level fixed effects. I cluster standard errors at the level of treatment, the district level. To choose the optimal bandwidth, I follow an MSE-optimal procedure (Calonico et al., 2014). I demonstrate robustness to other bandwidths. Following Gelman and Imbens (2019), I apply linear functions within the optimal bandwidth. I test for robustness to higher-order polynomials. The primary coefficient of interest is β_1. If the identification assumption is satisfied, the estimator can be interpreted as the local average treatment effect (LATE) of receiving the underbanked status for a district with a ratio equal to the cutoff.

B. Identification Assumption

The identification assumption of this setting is the continuity of all characteristics other than being underbanked at the cutoff. This assumption is violated if agents precisely manipulate the ratio of their district. Consider the following to understand how systematic differences could be introduced by manipulation. Assume local governments learn about the policy and want to benefit from more banks in their area. Also, assume they can manipulate the population-to-branch ratio, moving from just below the cutoff to just above it. If these districts have a particularly healthy population, I would confuse their characteristics with a treatment effect of the policy.

Manipulation of the population-to-branch ratio is unlikely due to its construction. First, the numerator contains population data from the 2001 Census. To manipulate this historical data, local governments would have to have anticipated the detailed policy rule years before its implementation. Second, the denominator is the sum of the individual decisions of all banks in the district. The total number of bank branches in the first quarter of 2006 is not determined by a specific bank or bank type alone, making manipulation unlikely. Also, banks directly report their number of branches to the RBI, leaving no room for an intermediary party to manipulate. I also test empirically for manipulation.

The first implication of manipulation refers to the density of the forcing variable. If local governments indeed manipulate their population-to-branch ratio, there should be more districts just above the cutoff than just below. At first glance, there is no evidence of this in Figure 3(a). To formally test for smoothness around the cutoff, I use the McCrary (2008) density test, depicted in Figure A3. I obtain an estimator of -0.1998 with a p-value of 0.8416, suggesting that I should not reject smoothness around the cutoff. The second implication of manipulation is that districts just above the cutoff should differ from those just below the cutoff before the policy. Assume, for example, that local governments that can manipulate their ratio have a healthier population. In this case, I would observe discontinuities in pre-policy health measures.

To test for smoothness before the policy, I utilize the RBI Master Office File (2004), the IHDS I (2004/2005), and the Economic Census (2005). Results are depicted in Table 1. Columns 1 and 2 show the mean for all treated and control observations. Columns 3 and 4 depict means only for observations within the optimal bandwidth. Column 5 reports the fuzzy RDD coefficients, referring to β_1 as defined above. As expected, all coefficients are statistically insignificant. Treatment districts do not have significantly more branch licenses or actual branches before the policy introduction. Households in treatment districts are not significantly healthier than those in control districts before the policy, neither for non-chronic nor for chronic illnesses. For non-chronic illnesses, I observe smoothness in the incidence of disease in the past 30 days, total days household members were ill, as well as days of work or school they missed due to an illness. For chronic illnesses, I observe smoothness in the incidence of disease and days of work or school they missed due to an illness. I also demonstrate smoothness for employment, households' financial access, and healthcare supply. Employment is smooth before the policy, households are not more likely to own financial products in treatment districts, and healthcare providers are not more likely to be financed mainly by a loan or have more presence in treatment districts. Correspondingly, I observe graphical smoothness in Figure

4. Additionally, I use the SHRUG data to show that village- and town-level general economic activity and population characteristics are smooth (Table A5). Taken together, these tests suggest that there was no manipulation.

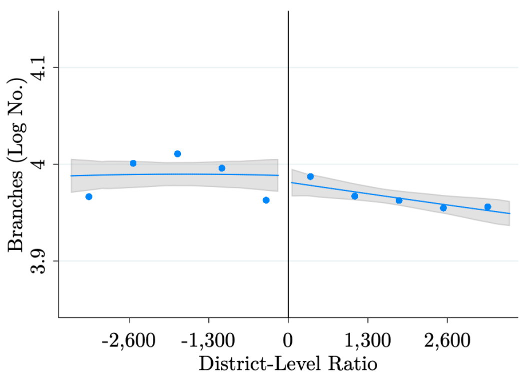
A second potential threat to identification is migration. If households migrate to treatment districts due to increased bank presence and these households are healthier, I would confuse their characteristics with a treatment effect of the policy. I have detailed data on migration that allows me to test for this threat. Less than 0.5 percent of households report that they moved to their current location from another district in the five years before the IHDS II (2011/2012). The coefficient on this migration pattern is insignificant when formally testing for it as described in the regression framework (Table A6).

Finally, I demonstrate that other policies do not threaten identification. The concern is that I may mistake discontinuities around the cutoff for the effect of the 2005 RBI policy when they stem from other policies. To my knowledge, no other policy uses the same cutoff rule described in this paper. For other nationwide policies to coincidentally threaten identification, they would need to be significantly more likely to be implemented in this study's treatment districts (Moscoe et al., 2015). Otherwise, their impact would be smooth around the cutoff. While many policies define certain priority districts, these are unlikely to be identical or highly correlated to treatment districts in this setting. The reason is that priority districts are often defined according to the target of the policy, for instance, certain health indicators. In Discussion A2, I describe other nationally implemented policies, including those issued by the Ministry of Health and Family Welfare, the Ministry of Women and Childhood Development, the Ministry of Labour and Employment, and other policies not directly related to health, such as the National Rural Employment Guarantee Act (NGREA), a labor guarantee program. For each policy, I collect a list of priority districts and map them to the 2001 Census borders. I then create an indicator that is one if a district is defined as a priority district under a specific policy and zero otherwise. Using this indicator variable as an outcome, I test whether the policy was significantly more likely to be implemented in treatment districts (Table A7). All coefficients are statistically insignificant. I provide further evidence on the distribution of priority districts in Table A8. Correlation coefficients between an indicator for priority district and an indicator for being underbanked within the bandwidth range from -0.08 to 0.22. This evidence suggests that other policies do not threaten causal identification. In summary, tests of the identification assumption strengthen the causal interpretation of my findings.

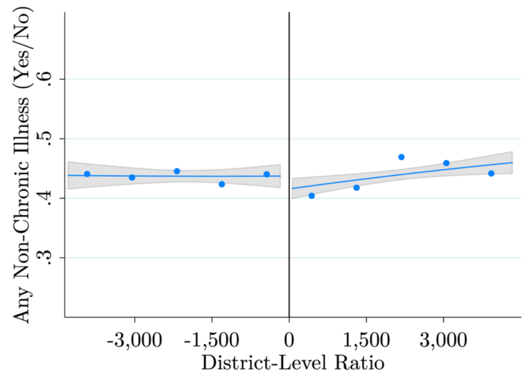
Table 1: Smooth Pre-Policy Covariates

	All observations		Within bandwidth		RDD
	Treated	Not treated	Treated	Not treated	Coefficient
	(1)	(2)	(3)	(4)	(5)
Banks					
Branch licenses (log no.)	4.04	4.74	3.90	4.13	0.02
	(0.80)	(0.88)	(0.89)	(1.05)	(0.02)
Branches (log no.)	4.02	4.74	3.72	4.27	0.01
	(0.81)	(0.88)	(0.95)	(1.03)	(0.02)
Health					
Non-chronic: any illness (yes/no)	0.52	0.39	0.47	0.40	-0.07
	(0.50)	(0.49)	(0.50)	(0.49)	(0.06)
Non-chronic: days ill (log no.)	1.13	0.79	0.95	0.82	-0.21
	(1.19)	(1.09)	(1.12)	(1.10)	(0.16)
Non-chronic: days missed (log no.)	0.78	0.55	0.60	0.62	-0.28
	(1.05)	(0.92)	(0.96)	(0.97)	(0.19)
Chronic: any illness (yes/no)	0.26	0.28	0.25	0.27	0.03
	(0.44)	(0.45)	(0.43)	(0.44)	(0.05)
Chronic: days missed (log no.)	0.68	0.70	0.62	0.70	-0.07
	(1.53)	(1.55)	(1.49)	(1.58)	(0.19)
Households' Financial Access					
Any loan (yes/no)	0.48	0.36	0.47	0.40	0.02
	(0.50)	(0.48)	(0.50)	(0.49)	(0.09)
Largest loan amt (log Rs)	4.43	3.56	4.41	3.86	0.39
	(4.70)	(4.83)	(4.77)	(4.84)	(0.78)
Largest loan from bank (yes/no)	0.11	0.13	0.12	0.12	-0.02
	(0.31)	(0.33)	(0.32)	(0.32)	(0.02)
Employment					
Employment (log no.)	11.63	11.79	11.41	11.58	0.01
	(0.87)	(1.01)	(1.10)	(1.26)	(0.12)
Healthcare Supply					
Institutional loan (share)	0.02	0.04	0.02	0.03	0.00
	(0.03)	(0.04)	(0.02)	(0.04)	(0.01)
Healthcare providers (log no.)	5.57	5.83	5.34	5.50	-0.15
	(0.97)	(1.14)	(1.22)	(1.39)	(0.16)

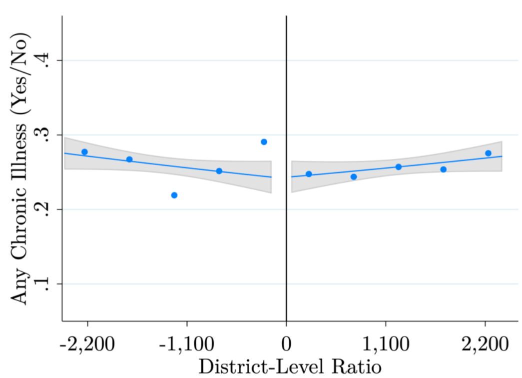
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are in parentheses and clustered on the district level. Bandwidth abbreviated by BW. Data RBI Master Office File (1998-2016), IHDS I (2004/2005), and Economic Census (2005). District and household level. Count and amount variables are transformed to log and winsorized at the 1st and 99th percentile. Variables depicted here are later used in post-policy regressions, explained in more detail in respective tables.



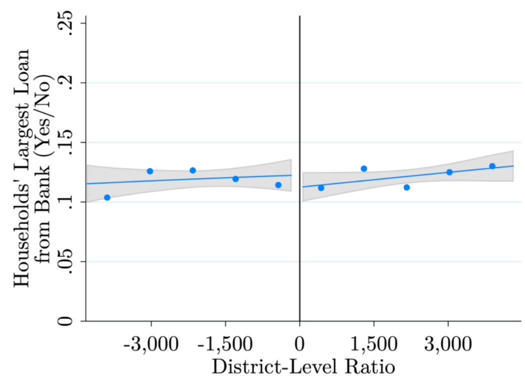
(a) Pre-Policy (RBI, 2004)



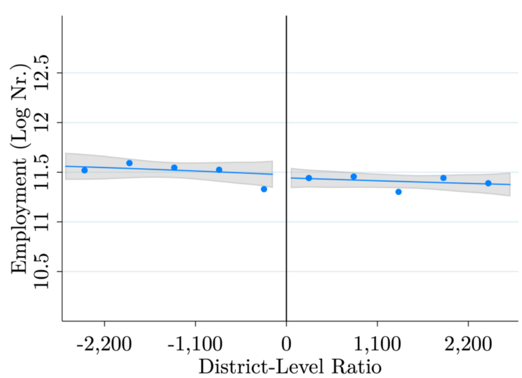
(b) Pre-Policy (IHDS, 2004/2005)



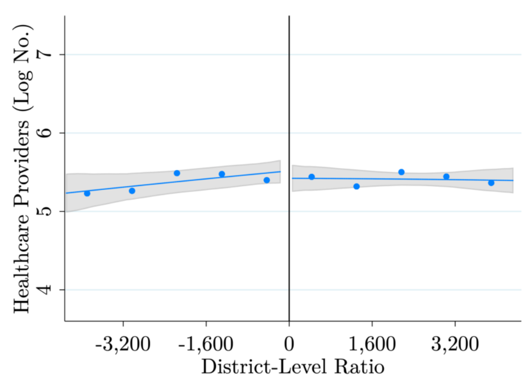
(c) Pre-Policy (IHDS, 2004/2005)



(d) Pre-Policy (IHDS, 2004/2005)



(e) Pre-Policy (Economic Census, 2005)



(f) Pre-Policy (Economic Census, 2005)

Figure 4. Smooth Pre-Policy Covariates. These graphs show binned means to the left and right of the cutoff within the optimal bandwidth. They also show local linear polynomials to the left and right of the cutoff, with 95 percent confidence intervals in gray. The cutoff is normalized to zero.

IV. Banks Open Branches

In the first step of the analysis, I provide evidence that the policy resulted in meaningful bank entry in treatment districts. I examine two outcomes: the number of branch licenses and the number of branches. Since I observe years between 1997 and 2016, I test both for smoothness pre-policy and discontinuities post-policy. In Table 2, I examine the number of branch licenses and branches in 2004, one year before the policy, and in 2010, five years after the policy.[1] As expected, coefficients in the year before the policy are statistically insignificant. Treatment districts have neither more branch licenses nor more branches than control districts. Post-policy, I observe statistically significant discontinuities in both branch licenses and branches. In 2010, treatment districts have 21 percent more branch licenses and 19 percent more branches than control districts (Figures 5(a) and 5(b)). The latter corresponds to an increase of 27 branches, compared to 142 branches in control districts. Private banks were the driving force behind this expansion, potentially boosting competition in the banking sector (Table A9). Private branches increase by 60%, while public branches increase by 12% relative to the control mean. In other words, while 17% of branches in the control are private, 53% of the new branches are private. The branch entry is economically meaningful. Utilizing administrative data on deposit accounts and credit amounts from the RBI, I find that

[1]Tables that describe treatment effects contain the following information: The first line provides the main coefficient of interest, β_1. This is followed by the control mean within the optimal bandwidth and the first stage coefficient, α_1. Following that are the optimal bandwidth and the number of observations within the optimal bandwidth. The next line, observations, describes the total size of the sample before conditioning on the bandwidth. Finally, the last line indicates whether the regression includes baseline controls.

Table 2: Banks Open Branches

	Pre-policy (2004)		Post-policy (2010)	
	Branch licenses (log no.) (1)	Branches (log no.) (2)	Branch licenses (log no.) (3)	Branches (log no.) (4)
Treated	0.02	0.01	0.19***	0.17***
	(0.02)	(0.02)	(0.05)	(0.06)
Control Mean	4.17	4.17	4.55	4.54
First Stage	0.81	0.80	0.80	0.80
Bandwidth	3,490	3,621	2,972	3,329
Obs. in BW	223	230	196	213
Observations	561	562	561	561
Baseline Control	Yes	Yes	Yes	Yes

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are in parentheses and clustered on the district level. Bandwidth abbreviated by BW. Data RBI Master Office File. District level. All variables are transformed into log form and winsorized at the 1st and 99th percentile. The variable from 1997 is included as a baseline control.

treatment districts have 11% more deposit accounts (161,977 new accounts) and 15% more credit (148 million USD) (Table A10). Credit by private banks increases by 54%, contributing 77 million USD to total credit gained. While private banks provide 15% of total credit in the control, they provide 52% of the new credit. To summarize, the policy introduced exogenous and economically meaningful bank entry.

Providing further support of design, the dynamics of the branch opening follow the policy timing (Figures 5(c) and 5(d)). As expected, there is smoothness around the cutoffs before the policy, and coefficients become statistically significant after the policy. The reaction in branch licenses issued is immediate: the coefficient on branch licenses becomes statistically significant in 2006 when the final list of under-banked districts is published. As expected, the branch reaction is slightly lagged by

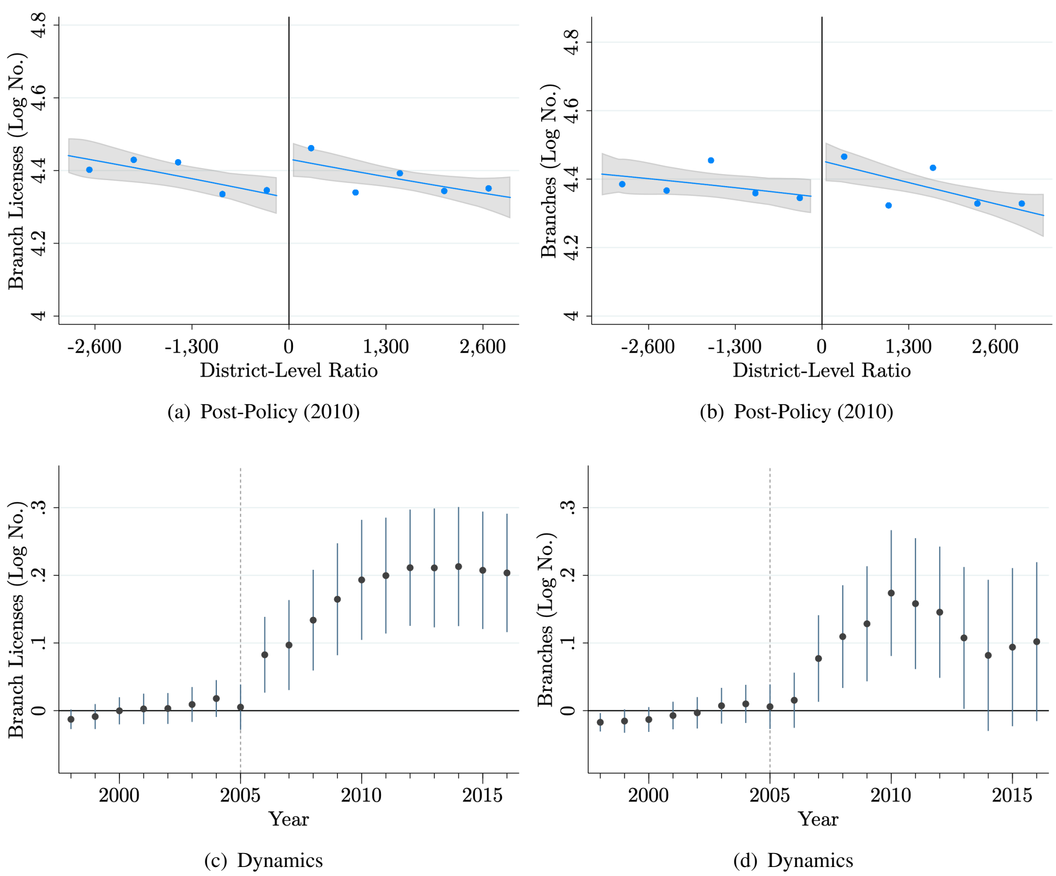


Figure 5. RBI Issues Licenses and Banks Open Branches. Figures 5(a) and 5(b) depict the discontinuities in branch licenses and branches five years after the policy was introduced. Respective regressions are described in Columns 3 and 4 of Table 2. Figures 5(c) and 5(d) depict the dynamic effects of branch licenses and branches.

one year: the coefficient becomes statistically significant in 2007. There is another pattern that the policy can explain. In 2010, as discussed in Section I, the RBI allowed banks to open branches without licenses in eight states. The observed pattern in the dynamics—a stagnation in the coefficient on licenses issued and a decrease in the coefficient on the number of branches—corresponds exactly to what one expects to see if banks increasingly open branches in districts to the left of the cutoff (remaining in the control group) in states where licenses are not required. While the change in the policy attenuates the difference in branches between treatment and control districts after 2010, treatment districts have historically been exposed to more branches. One can conclude that the branch opening dynamics follow the RBI policy.

Standard robustness and placebo tests on bank outcomes are discussed in Section VIII, but one placebo test that emerges from the design of the policy is outlined here. One type of bank, regional rural banks, is excluded from the policy. Consequently, one does not expect to observe positive coefficients for this bank type. I test for discontinuities in branch licenses and branches of regional rural banks in 2010 (Table A11), and coefficients in the placebo test are insignificant.

One question the reader may remain curious about is whether these new branches are profitable for the banks. Answering this question requires data on branch profits. Unfortunately, neither the RBI nor any other institution provides this data. Without data on branch profitability, it is not possible to estimate the costs of the policy, which are potentially carried by the financial sector. This paper does not target a full policy evaluation but instead uses the policy to obtain exogenous variation in bank presence. However, it is possible to make one specific statement on profitability: As banks indeed react to the policy, the combination of opening a branch in an underbanked district and obtaining a license for another location appears profitable for banks.

V. Non-Chronic Diseases Improve

Can bank presence move the needle on households' health? To answer this question, I next turn to households' health status. Consistent with other health economics studies, I investigate two primary outcomes: non-chronic and chronic illnesses. Non-chronic illnesses are frequent in many developing countries; they include illnesses such as fever, diarrhea, and cough. In the IHDS II, 55% of households experience a non-chronic disease in a given month. Conditional on the illness, households face ten days of illness, summed over the members. They spend 621 rupees (6% of total monthly consumption) and lose six days of work or school. Thus, non-chronic illnesses are frequent and consequential. To test for changes in health status, I rely on two nationally representative household surveys. The IHDS II (2011/2012) allows me to measure non-chronic and chronic diseases six years after the policy introduction. The DHS (2015/2016) allows me to replicate the results in the long term, ten years later.

The IHDS II (2011/2012) provides information on whether any household member was ill in the past 30 days with a non-chronic disease, which refers to fever, diarrhea, or cough (Table 3, Column 1). Additionally, I observe the number of days household members were ill (Column 2) or could follow usual activities such as work or school (Column 3), aggregated over members. I find improvements for non-chronic illnesses. Households in treatment districts are 19 percentage points less likely to have a member who suffered from a non-chronic disease in the past month. Comparing this to the control mean corresponds to a decrease of 36%. The discontinuity is depicted in Figure 6(a). As households gain, on average, two healthy days, they also increase their labor supply and school attendance. While I observe a positive impact on non-chronic illnesses, I do not find an improvement in chronic illnesses like heart disease or cancer (Table 3, Columns 4 and 5). This could have multiple reasons. First, unlike non-chronic diseases, the prevalence of chronic diseases is likely much less responsive to household investments in food and sanitation. Additionally, even with an increase in healthcare demand or supply, healthcare providers might not be equipped to deal with these diseases as they lack expertise or expensive equipment.

Table 3: Fewer Non-Chronic Illnesses

	Non-chronic illness			Chronic illness	
	Any (yes/no) (1)	Days ill (log no.) (2)	Days missed (log no.) (3)	Any (yes/no) (4)	Days missed (log no.) (5)
Treated	-0.19**	-0.43**	-0.61***	-0.00	-0.31
	(0.08)	(0.19)	(0.20)	(0.05)	(0.20)
Control Mean	0.52	1.08	0.79	0.39	1.04
First Stage	0.65	0.66	0.67	0.65	0.63
Bandwidth	2,204	2,312	2,440	2,189	2,087
Obs. in BW	11,986	12,927	13,595	11,953	10,518
Observations	36,673	38,375	38,485	36,673	36,673
Baseline Control	No	No	No	No	No

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are in parentheses and clustered on the district level. Bandwidth abbreviated by BW. Data IHDS II (2011/2012). Household level. All variables measured in days are in log and winsorized at the 1st and 99th percentile. Non-chronic illnesses include fever, diarrhea, and cough in the past 30 days. Chronic illnesses include, for instance, heart disease and cancer, ever diagnosed (column 4) or days unable to work in the past 12 months (column 5).

How does the effect size on non-chronic illnesses compare to other health interventions? Table A12 provides an overview of meta-studies and other benchmark papers, showing that the effect size is in the middle of the range of other successful health interventions in developing countries. The health economics literature contextualizes these effect sizes. For many non-chronic diseases, there exist highly effective and relatively cheap treatments, for example, oral rehydration solutions for diarrhea (Banerjee and Duflo, 2011; Dupas and Miguel, 2017). Additionally, improving the health of some households could have spillover effects, reducing infection risks of others (Kremer and Glennerster, 2011). Thus, the effect sizes on non-chronic diseases are comparable to the literature and sensible given the context. To provide further confidence in the effect, I show that outcomes in Table 3 are smooth on baseline (see Figure 4(b), Table 1, and Table A13) and robust to controlling for baseline measures (see Table A14). Table A15 discusses robustness to different transformations, including level and inverse hyperbolic sine. Further robustness is discussed in Section VIII.

To replicate my findings from the IHDS and obtain long-term effects, I utilize the DHS (2015/2016) in Table 4. I find that households are six percentage points less likely to have a child with fever, diarrhea, or cough in the past two weeks. Since in the control group, around every fourth household has an ill child, this corresponds to a mean change of 23%. Non-chronic diseases are collected for children below the age of five. I use the proxy of healthcare visits to understand diseases for other family members. Visits are a function of health status, demand, and sup-

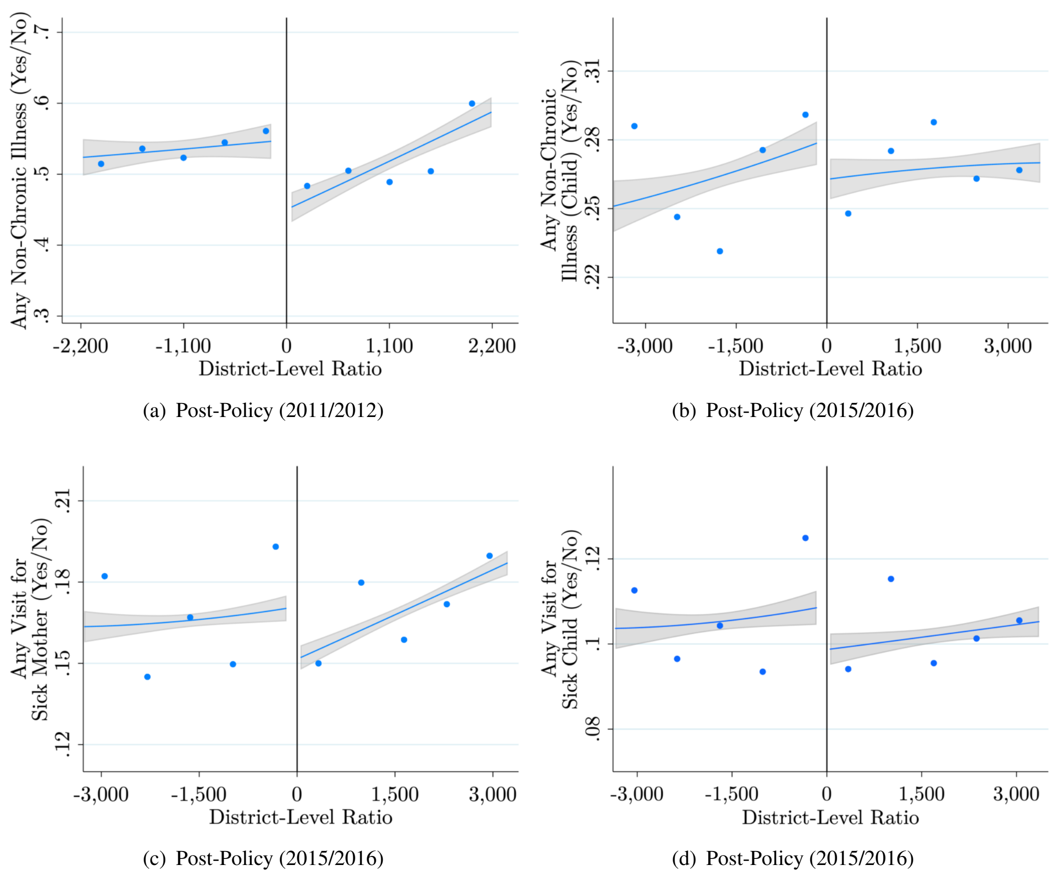
Table 4: Results Hold in Second Survey

	Non-chronic illness	Any illness	
	Sick child (yes/no) (1)	Visit for sick child (yes/no) (2)	Visit for sick mother (yes/no) (3)
Treated	-0.06* (0.03)	-0.02* (0.01)	-0.05* (0.03)
Control Mean	0.27	0.11	0.17
First Stage	0.70	0.73	0.72
Bandwidth	3,539	3,383	3,277
Obs. in BW	66,658	187,208	182,318
Observations	171,471	471,985	471,985
Baseline Control	No	No	No

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are in parentheses and clustered on the district level. Bandwidth abbreviated by BW. Data DHS (2015/2016). Household level. Column 1 shows whether a child had a non-chronic disease (fever, diarrhea, cough) in the past two weeks. Columns 2 and 3 indicate healthcare visits for any illness in the past three months. The table refers to children below five. Data is missing for households without children below five in (1) and without eligible women in Columns (2) and (3).

ply; thus, they do not perfectly reflect the incidence of illnesses. With this caveat, results are consistent with households getting healthier. They are two percentage points less likely to go to a healthcare provider for treatment of a sick child and five percentage points less likely for treatment of a sick mother. Discontinuities are depicted in Figure 6. Note that visits are measured ten years after the policy. It is likely that healthcare visits increased in the first years after the policy, but as households get healthier, they require these services less. With positive effects in the DHS, two different surveys indicate that bank presence improves non-chronic diseases.

The reader may ask whether the results are biased by differential reporting on health status in treatment and control. First, the bias would go in the opposite direction. If banks positively affect households' awareness about diseases or the



(a) Post-Policy (2011/2012)

(b) Post-Policy (2015/2016)

(c) Post-Policy (2015/2016)

(d) Post-Policy (2015/2016)

Figure 6. Health Improves. These graphs show binned means to the left and right of the cutoff within the optimal bandwidth. They also show local linear polynomials to the left and right of the cutoff, with 95 percent confidence intervals in gray.

likelihood of being diagnosed, this would create an upward bias in the estimate, reducing the chance of detecting a decrease in reported diseases. Second, I can study outcomes that are not self-reported, such as vaccinations reported on a vaccination card. As expected, I find positive effects for these outcomes (Table A16). Thus, self-reporting biases are unlikely to play a role.

VI. Banking Activities

After establishing the causal effect of bank presence on health, I turn towards providing suggestive evidence on how bank presence affects health. I take two steps of analysis. First, I investigate what specific banking activities are at play. Second, I examine how these banking activities translate into changes in households' health investments. Starting with banking activities, previous literature has demonstrated that banks extend credit to firms, thereby fostering economic activity and employment (Bruhn and Love, 2014; Fonseca and Matray, 2022). In addition to simulating employment that allows households to invest more in health, banks may impact health through three distinct activities. First, banks may offer savings accounts to households. Second, they might provide personal bank loans to households. Both savings accounts and bank loans could allow households to invest more in health when necessary. Finally, banks could extend credit to healthcare providers, allowing them to increase healthcare supply, an important factor for health status (Das and Hammer, 2005, 2014). I empirically test whether each of these activities is at play. Readers curious about whether one specific banking activity in isolation would have been sufficient to make an impact can turn towards randomized controlled trials. This paper instead takes a comprehensive view, relaxing the financial constraints of multiple actors in the economy simultaneously. This acknowledges the complexities of healthcare markets, where banks can stimulate both demand- and supply-side. Furthermore, it informs branch policies implemented globally, including in countries like Brazil and China.

As outlined in Section IV, total credit in the economy increases by 15%, or around 148 million USD in 2010. To test whether this credit stimulated an employment effect, I utilize the Economic Census pre-policy (2005) and post-policy (2013). I find that total employment increases in the economy by 12%. This effect size is consistent with other branch expansion policies focused on labor market outcomes (Bruhn and Love, 2014). The result is primarily driven by increased employment in the service sector. Employment is smooth pre-policy, as outlined in Table A18. Additionally, I test another hypothesis: the increase in business activity post-policy could have increased local tax revenue and, thus, government spending on health. However, empirically, I do not find any effects on government spending

Table 5: Employment Increases

	Total employment		
	Total (log nr.) (1)	Manufacturing (log nr.) (2)	Services (log nr.) (3)
Treated	0.12*	0.09	0.11*
	(0.07)	(0.10)	(0.06)
Control Mean	11.83	10.35	11.50
First Stage	0.80	0.78	0.80
Bandwidth	3,548	2,626	3,796
Obs. in BW	225	169	239
Observations	555	555	556
Baseline Control	Yes	Yes	Yes

* p < 0.1, ** p < 0.05, *** p < 0.01. Standard errors are in parentheses and clustered on the district level. Bandwidth abbreviated by BW. Data Economic Census (2013), aggregated by the SHRUG, which separates manufacturing and service employment. District level. Variables in log and winsorized at the 1st and 99th percentile.

on health-related categories (Table A19). This is consistent with the difficulties of local governments in collecting taxes in this context.

To explore the financial access of households, I utilize the IHDS (2011/2012). Households are asked whether they had any savings account or bank loan in the past five years. Savings accounts and bank loans could allow households to smooth consumption, thereby having funds available to invest in health when necessary. I find that households are significantly more likely to own a savings account. Households in treatment districts are 19 percentage points more likely to own a savings account compared to a control mean of 51 percent. In contrast, the average household in my sample is not more likely to have a bank loan. This is in line with studies that discuss low formal credit take-up in developing countries (Banerjee et al., 2015a;

Table 6: More Savings Accounts But Not Bank Loans for Households

	Savings account (yes/no) (1)	Bank loan (yes/no) (2)
Treated	0.19**	0.04
	(0.10)	(0.05)
Control Mean	0.51	0.23
First Stage	0.69	0.66
Bandwidth	3,023	2,370
Obs. in BW	16,911	13,093
Observations	38,348	38,351
Baseline Control	No	No

* p < 0.1, ** p < 0.05, *** p < 0.01. Standard errors are in parentheses and clustered on the district level. Bandwidth abbreviated by BW. Data IHDS II (2011/2012). Household level. Households are asked whether they had any savings account or bank loan in the past five years.

Badarinza et al., 2019). Note that households might still take up informal debt to cope with medical expenses (Ramadorai, 2017). Take-up of financial instruments by households is balanced pre-policy (Table A17). Not all outcome variables are available pre-policy, in which case similar dimensions of financial access are shown to be smooth. Thus, these results provide suggestive evidence that savings accounts to households played a role in improving health, while households' formal medical debt is unlikely to play a role.

Finally, I explore the relationship between banks and healthcare providers. Households in developing countries have limited access to good healthcare services, an important determinant of health status (Banerjee et al., 2004). Many households are highly unsatisfied with the sector. Thirty-six percent of households in the DHS (2015/2016) state that distance to the closest health facility is a big problem. Fifty-two percent report that personnel absenteeism is a big issue, and 53 percent have large problems with drug availability at healthcare facilities. If bank presence allows healthcare providers to relax their credit constraints, this could allow investing in new healthcare facilities, providing monetary incentives for medical personnel to decrease absenteeism rates, or purchasing drugs on stock. For bank presence to increase supply, two conditions must be satisfied: healthcare providers generally rely on bank loans and are credit-constrained.

To examine whether healthcare providers rely on bank loans, I use two datasets: the Prowess database and the Economic Census. The Prowess provides detailed financial information about a sample of relatively large healthcare providers from 1988 to 2017. These observations are limited to only 89 districts; thus, I use this data only for descriptive purposes. I examine statistics for the 385 companies in the Prowess data that conduct hospital activities, averaging over the years present in the database. These companies have a broad asset range of USD 2,000 to 410 million, with a mean asset size of 15 million. Of these companies, 65 percent have a bank loan. For those with a bank loan, the mean size of the loan is USD 5.09 million, corresponding to 33 percent of their total assets. Bank loans as a financial instrument are used by companies across the size range (see Figure A4). Thus, relatively large healthcare providers rely on bank loans.

To examine whether smaller healthcare providers also rely on bank loans, I turn to the Economic Census, which only collects data on the major source of finance. It does not contain additional balance sheet data. Institutional loans are rarely the major source of finance for healthcare providers: only 1.59 percent of establishments with hospital activities list loans as their major source of finance. Instead, commonly cited major sources of finance are self-finance with 44 percent and government sources with 39 percent. The fact that few healthcare providers cite institutional loans as their major source of finance does not imply that they do not rely on bank loans. Healthcare providers are only slightly less likely to cite an institutional loan as their major source of finance than all businesses (2.11 percent). This provides cautious evidence that they rely on bank loans. Finally, the question arises whether the second condition is satisfied: that healthcare providers are credit-constrained. While there is no evidence available for healthcare providers specifically, academic research has established that, generally, medium-sized companies (Banerjee and Duflo, 2014) and small businesses (de Mel et al., 2008) in developing countries are credit constrained. The conditions are met for credit access to healthcare providers to play a role.

Finally, I investigate the Economic Census to learn how healthcare activity responds to bank presence (Table 7). I find that treatment districts have a one percentage point or 65% increase in the share of healthcare providers primarily financed by a loan. As outlined, this likely masks an overall larger effect in absolute terms as the survey only asks about the biggest loan and not any loan. In equilibrium, I observe an increase in the number of healthcare providers. These are not large healthcare providers; they have, on average, only seven employees. Table A20 shows the smoothness of financial access and healthcare supply before the policy. Consistent with the increase in healthcare supply post-policy, I find in the DHS that households are more likely to shift to private providers (Table A21). To summarize, bank presence likely affects health through established activities (firm credit, household savings accounts) and understudied aspects (credit to healthcare providers) but not through personal bank loans.

Table 7: Healthcare Supply Increases

	Healthcare Providers	
	Institutional loan (share) (1)	Number (log nr.) (2)
Treated	0.01**	0.89***
	(0.00)	(0.33)
Control Mean	0.01	5.96
First Stage	0.79	0.80
Bandwidth	2,435	3,127
Obs. in BW	164	204
Observations	556	556
Baseline Control	Yes	Yes

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are in parentheses and clustered on the district level. Bandwidth abbreviated by BW. Data Economic Census (2013). District level. Variables in log and winsorized at the 1st and 99th percentile.

VII. Households' Health Investments

How exactly did these banking activities translate into health improvements? To investigate this further, I study households' health-related investments. Households might improve their health by spending more on low-fixed-cost items such as food or hygiene, high-fixed-cost items such as toilet facilities or fridges, or by increasing their healthcare demand. Consistent with a gradual increase in availability of resources through employment and savings accounts, I find that households in treatment districts consume more meals and spend more on hygiene expenses (Table A22). Hygiene expenses include soap, insecticides, and toilet articles. Both higher food consumption by strengthening the body and higher hygiene expenses by reducing infection risks can positively impact non-chronic diseases. In alignment with a lack of credit take-up, I do not find evidence that households invest in high-fixed-cost items such as toilet facilities or fridges (Table A22). Finally, I consider households' healthcare demand. One might be tempted to take healthcare expenditure or visits for diseases as proxies for demand. In the context of this study, they are not suitable proxies. To understand why, note that these variables are measured six to ten years after the policy introduction. It is possible that households spent more and visited more in the years after the policy introduction, are healthier at the point of the survey, and require respective healthcare services less. A negative effect on spending and visits then does not reflect a decrease in healthcare demand but an increase in health status. Thus, I explore an alternative proxy for healthcare demand. I examine healthcare utilization of services that should not decrease as households get healthier. In particular, I consider vaccinations and pregnancy care. Both are higher in treatment districts, providing suggestive evidence that households increase their healthcare demand (Table A23). To summarize, banking activities likely improved health by increasing households' spending on low-fixed-cost items and healthcare demand, but not through households' investments in high-fixed-cost items.

VIII. Robustness and Placebo Tests

To demonstrate the robustness of my results, I initially test whether coefficients remain statistically significant for different bandwidth choices. I examine bandwidth multipliers in the range of 0.50 to 2.00, in steps of 0.25. For instance, if the MSE-optimal bandwidth (Calonico et al., 2014) is 2,000, I examine bandwidths from 1,000 to 4,000. Results are described in Table A24 as well as Figures A6 and A7. Considering the optimal bandwidth with multipliers of 0.75 and 1.25, 73 percent remain statistically significant. Examining bandwidth multipliers of 0.50 and 1.50, 59 percent remain statistically significant. This suggests that results are robust to different bandwidth multipliers.

In a second approach, I examine different bandwidth selectors. The default is an MSE-optimal bandwidth selector by Calonico et al. (2014) that chooses identical bandwidths to the left and to the right of the cutoff. In Tables A25 and A26, I also consider an MSE-optimal selector that separately chooses bandwidths to the left and to the right of the cutoff. Additionally, I examine a selector by Calonico et al. (2020) that optimizes the coverage error rate (CER). I again consider the selector with identical and different bandwidths to the left and right of the cutoff. Figure A8 summarizes the results. I find that 72 to 82 percent of results remain statistically significant. This suggests that results are robust to different bandwidth selectors.

Results are also robust considering possible bias corrections due to the MSE-optimal bandwidth selector, discussed by Calonico et al. (2014) and Cattaneo and Vazquez-Bare (2017). This is depicted in Tables A25 and A26, and summarized in Figure A12. All of the coefficients remain statistically significant, suggesting that findings are highly robust to these adjustments.

I next examine robustness with respect to polynomial degrees. Gelman and Imbens (2019) argue that researchers should apply linear or quadratic approximations. Additionally, I examine the robustness of polynomials of degree three. Findings are described in Table A27 and summarized in Figure A9. For polynomials of degree two, 91 percent of outcomes remain statistically significant. For polynomials of degree three, I find that 72 percent of outcomes remain significant. In summary, results are highly robust to alternative polynomials.

Another classical regression discontinuity robustness test is to examine smoothness around placebo cutoffs. I examine three placebo cutoffs on each side of the normalized true cutoff (zero): $\pm$ 1,000, 2,000, and 3,000. This choice of placebo cutoffs ensures enough observations around the placebo cutoff to conduct an analysis. Evidence is provided in Table A28 and summarized in Figure A10. On average, 12 percent of outcomes are statistically significant. Thus, I find little evidence of discontinuities at placebo cutoffs.

Finally, I test whether results are robust to adjustments for multiple hypothesis testing and spatial correlation of standard errors in Table A29, summarized in Figure A11. To address concerns regarding multiple hypothesis testing, I adjust for the false discovery rate, following Anderson (2008). The false discovery rate is the expected proportion of rejections that are Type I errors (false rejections). To adjust for spatial correlation of standard errors, I adjust for Conley standard errors (Conley, 1999) in district-level regressions. Since the most granular location data available for households is their district, I do not adjust household-level regressions for spatial correlation. Results are robust to multiple hypothesis testing and spatial correction. In summary, the standard robustness and placebo tests support the validity of the findings.

IX. Conclusion

What are the general equilibrium effects of banks? While previous work has focused on labor market effects, this study demonstrates that banks can contribute towards tackling hard-to-crack development challenges, focusing on the third UN Sustainable Development Goal of improving health. This paper utilizes a 2005 RBI policy to obtain exogenous variation in bank presence, applying a regression discontinuity design. After establishing that the policy introduced exogenous and economically meaningful bank entry, I examine two nationally representative household-level surveys six and ten years after the policy introduction. Both surveys confirm that banks can move the needle on non-chronic diseases. There is no effect on chronic illnesses. The paper finds suggestive evidence for three out of four banking activities. Banks stimulate employment, offer savings accounts to households, and credit to healthcare providers. I do not find evidence of the personal medical bank debt narrative. For households, these banking activities translate into higher spending on food and hygiene as well as increased healthcare demand. I do not find evidence that households' investments in high-fixed-cost items play a role, consistent with a lack of credit take-up.

This paper has important implications for policy and future research. Policymakers can conclude that incentivizing banks to enter underserved locations can benefit their citizens' health. They might also focus on the interaction of banks with local providers of services they want to foster. Indeed, the RBI announced a policy in May 2021 to incentivize banks to quickly deliver credit to healthcare providers in light of the COVID crisis, announcing plans to inject USD 6.78 billion of liquidity. This paper also speaks to researchers. The study encourages further exploration into the impact of finance on various dimensions of well-being, including education. Gaining insights into these inquiries can substantially advance our understanding of the impact of banks on households.

References

Adelino, Manuel, Katharina Lewellen, and Anant Sundaram, "Investment decisions of nonprofit firms: Evidence from hospitals," *The Journal of Finance*, 2015, *70* (4), 1583–1628.

_ , _ , and W Ben McCartney, "Hospital financial health and clinical choices: evidence from the financial crisis," *Management Science*, 2022, *68* (3), 2098–2119.

Agarwal, Sumit, Shashwat Alok, Pulak Ghosh, Soumya Ghosh, Tomasz Piskorski, and Amit Seru, "Banking the unbanked: What do 255 million new bank accounts reveal about financial access?," Columbia Business School Research Paper 17-12 2017.

Aghamolla, Cyrus, Pinar Karaca-Mandic, Xuelin Li, and Richard T Thakor, "Merchants of death: The effect of credit supply shocks on hospital outcomes," *Available at SSRN 4621779*, 2023.

Andersen, Asger Lau, Rajkamal Iyer, Niels Johannesen, Mia Jørgensen, and José-Luis Peydró, "Household leverage and mental health fragility," 2022.

Anderson, Michael L, "Multiple inference and gender differences in the effects of early intervention: A reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects," *Journal of the American statistical Association*, 2008, *103* (484), 1481–1495.

Angelucci, Manuela, Dean Karlan, and Jonathan Zinman, "Microcredit impacts: Evidence from a randomized microcredit program placement experiment by Compartamos Banco," *American Economic Journal: Applied Economics*, 2015, *7* (1), 151–82.

Antill, Samuel, Jessica Bai, Ashvin Gandhi, and Adrienne Sabety, "Healthcare provider bankruptcies," *Working paper*, 2023.

Asher, Sam and Paul Novosad, "Socioeconomic High-resolution Rural-Urban Geographic Dataset for India (SHRUG)," Harvard Dataverse, https://doi.org/10.7910/DVN/DPESAK, 2019. (Accessed January 5, 2020).

_ , Tobias Lunt, Ryu Matsuura, and Paul Novosad, "Development research at high geographic resolution: An analysis of night lights, firms, and poverty in

India using the SHRUG open data platform," 2021. World Bank Group Policy Research Working Paper 9540.

Attanasio, Orazio, Britta Augsburg, Ralph De Haas, Emla Fitzsimons, and Heike Harmgart, "The impacts of microfinance: Evidence from joint-liability lending in Mongolia," *American Economic Journal: Applied Economics*, 2015, *7* (1), 90–122.

Augsburg, Britta, Ralph De Haas, Heike Harmgart, and Costas Meghir, "The impacts of microcredit: Evidence from Bosnia and Herzegovina," *American Economic Journal: Applied Economics*, 2015, *7* (1), 183–203.

Aydin, Deniz, "Consumption response to credit expansions: Evidence from experimental assignment of 45,307 credit lines," *American Economic Review*, 2022, *112* (1), 1–40.

Bachas, Pierre, Paul Gertler, Sean Higgins, and Enrique Seira, "How debit cards enable the poor to save more," *The Journal of finance*, 2021, *76* (4), 1913–1957.

Badarinza, Cristian, Vimal Balasubramaniam, and Tarun Ramadorai, "The household finance landscape in emerging economies," *Annual Review of Financial Economics*, 2019, *11*, 109–129.

Banerjee, Abhijit and Esther Duflo, *Poor Economics*, New York: PublicAffairs, 2011.

_ , Angus Deaton, and Esther Duflo, "Wealth, health, and health services in rural Rajasthan," *American Economic Review*, 2004, *94* (2), 326–330.

_ , Dean Karlan, and Jonathan Zinman, "Six randomized evaluations of microcredit: Introduction and further steps," *American Economic Journal: Applied Economics*, 2015, *7* (1), 1–21.

_ , Emily Breza, Esther Duflo, and Cynthia Kinnan, "Can microfinance unlock a poverty trap for some entrepreneurs?," Technical Report, National Bureau of Economic Research 2019.

_ , Esther Duflo, Rachel Glennerster, and Cynthia Kinnan, "The miracle of microfinance? Evidence from a randomized evaluation," *American Economic Journal: Applied Economics*, 2015, *7* (1), 22–53.

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

Bang, Abhay T, Rani A Bang, Sanjay B Baitule, M Hanimi Reddy, and Mahesh D Deshmukh, "Effect of home-based neonatal care and management of sepsis on neonatal mortality: field trial in rural India," *The lancet*, 1999, *354* (9194), 1955–1961.

Barboni, Giorgia, Erica Field, and Rohini Pande, "Rural banks can reduce poverty: evidence from 870 Indian villages," Working Paper 2021.

Beaman, Lori, Dean Karlan, Bram Thuysbaert, and Christopher Udry, "Self-Selection into Credit Markets: Evidence from Agriculture in Mali," Working Paper 20387, National Bureau of Economic Research August 2014.

Björkman, Martina and Jakob Svensson, "Power to the people: evidence from a randomized field experiment on community-based monitoring in Uganda," *The Quarterly Journal of Economics*, 2009, *124* (2), 735–769.

Björkman-Nykvist, M, G Andrea, J Svensson, and D Yanagizawa-Drott, "Evaluating the impact of the Living Goods entrepreneurial model of community health delivery in Uganda: A cluster-randomized controlled trial," Mimeo 2014.

Breza, Emily and Cynthia Kinnan, "Measuring the equilibrium impacts of credit: Evidence from the Indian microfinance crisis," *The Quarterly Journal of Economics*, 2021, *136* (3), 1447–1497.

Brown, James R, J Anthony Cookson, and Rawley Z Heimer, "Growing up without finance," *Journal of Financial Economics*, 2019, *134* (3), 591–616.

Bruhn, Miriam and Inessa Love, "The real impact of improved access to finance: Evidence from Mexico," *The Journal of Finance*, 2014, *69* (3), 1347–1376.

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

Calonico, Sebastian, Matias D Cattaneo, and Max H Farrell, "Optimal bandwidth choice for robust bias-corrected inference in regression discontinuity designs," *The Econometrics Journal*, 2020, *23* (2), 192–210.

_ , _ , and Rocio Titiunik, "Robust nonparametric confidence intervals for regression-discontinuity designs," *Econometrica*, 2014, *82* (6), 2295–2326.

Cattaneo, Matias D and Gonzalo Vazquez-Bare, "The choice of neighborhood in regression discontinuity designs," *Observational Studies*, 2017, *3* (2), 134–146.

Célerier, Claire and Adrien Matray, "Bank-branch supply, financial inclusion, and wealth accumulation," *The Review of Financial Studies*, 2019, *32* (12), 4767–4809.

CMIE, "CMIE Prowess dx, Vintage March 2020," https://prowessdx.cmie.com/, 2020. (Accessed April 17, 2020).

Conley, Timothy G, "GMM estimation with cross sectional dependence," *Journal of Econometrics*, 1999, *92* (1), 1–45.

Crépon, Bruno, Florencia Devoto, Esther Duflo, and William Parienté, "Estimating the impact of microcredit on those who take it up: Evidence from a randomized experiment in Morocco," *American Economic Journal: Applied Economics*, 2015, *7* (1), 123–50.

CSO and MOSPI, "Economic Census, 2005," Central Statistical Office (CSO) and Ministry of Statistics Programme Implementation (MOSPI), http://microdata.gov.in/nada43/index.php/catalog/46, 2018. (Accessed April 27, 2019).

_ and _ , "Economic Census, 2013-2014," Central Statistical Office (CSO) and Ministry of Statistics Programme Implementation (MOSPI), http://microdata.gov.in/nada43/index.php/catalog/47, 2018. (Accessed April 27, 2019).

Das, Jishnu and Jeffrey Hammer, "Which doctor? Combining vignettes and item response to measure clinical competence," *Journal of Development Economics*, 2005, *78* (2), 348–383.

_ and _ , "Quality of primary care in low-income countries: facts and economics," *Annual Review of Economics*, 2014, *6* (1), 525–553.

de Mel, Suresh, David McKenzie, and Christopher Woodruff, "Returns to capital in microenterprises: evidence from a field experiment," *The Quarterly Journal of Economics*, 2008, *123* (4), 1329–1372.

Desai, Sonalde and Reeve Vanneman, "India Human Development Survey (IHDS), 2005," National Council of Applied Economic Research, New Delhi, Inter-University Consortium for Political and Social Research, https://doi.org/10.3886/ICPSR22626.v12, 2018. (Accessed February 2, 2019).

_ and _ , "India Human Development Survey-II (IHDS-II), 2011-12," Inter-University Consortium for Political and Social Research, https://doi.org/10.3886/ICPSR36151.v6, 2018. (Accessed February 2, 2019).

Doornik, Bernardus Ferdinandus Nazar Van, Armando R Gomes, David Schoenherr, and Janis Skrastins, "Financial access and labor market outcomes: Evidence from credit lotteries," *Available at SSRN 3800020*, 2021.

Dupas, Pascaline and Edward Miguel, "Impacts and determinants of health levels in low-income countries," in Abhijit Vinayak Banerjee and Esther Duflo, eds., *Handbook of Economic Field Experiments*, Vol. 2, Amsterdam: Elsevier, 2017, pp. 3–93.

_ **and Jonathan Robinson**, "Savings constraints and microenterprise development: Evidence from a field experiment in Kenya," *American Economic Journal: Applied Economics*, 2013, *5* (1), 163–92.

_ **, Dean Karlan, Jonathan Robinson, and Diego Ubfal**, "Banking the unbanked? Evidence from three countries," *American Economic Journal: Applied Economics*, 2018, *10* (2), 257–97.

Egger, Dennis, Johannes Haushofer, Edward Miguel, Paul Niehaus, and Michael W Walker, "General equilibrium effects of cash transfers: experimental evidence from Kenya," Technical Report, National Bureau of Economic Research 2019.

Fiorin, Stefano, Joseph Hall, and Martin Kanz, *How Do Borrowers Respond to a Debt Moratorium?: Experimental Evidence from Consumer Loans in India*, Centre for Economic Policy Research, 2023.

Fonseca, Julia and Adrien Matray, "Financial Inclusion, Economic Development, and Inequality: Evidence from Brazil," Technical Report 2022.

_ **and Bernardus Van Doornik**, "Financial development and labor market outcomes: Evidence from Brazil," *Journal of Financial Economics*, 2022, *143* (1), 550–568.

Garber, Gabriel, Atif R Mian, Jacopo Ponticelli, and Amir Sufi, "Consumption Smoothing or Consumption Binging? The effects of government-led consumer credit expansion in Brazil," Technical Report, National Bureau of Economic Research 2021.

Gelman, Andrew and Guido Imbens, "Why high-order polynomials should not be used in regression discontinuity designs," *Journal of Business & Economic Statistics*, 2019, *37* (3), 447–456.

Gertler, Paul, "Do conditional cash transfers improve child health? Evidence from PROGRESA's control randomized experiment," *American Economic Review*, 2004, *94* (2), 336–341.

Ghosh, Pulak and Nishant Vats, "Safety Nets, Credit, and Investment: Evidence from a Guaranteed Income Program," *Credit, and Investment: Evidence from a Guaranteed Income Program (November 1, 2022)*, 2022.

Giné, Xavier and Martin Kanz, "The economic effects of a borrower bailout: evidence from an emerging market," *The Review of Financial Studies*, 2018, *31* (5), 1752–1783.

Gupta, Atul, Sabrina T Howell, Constantine Yannelis, and Abhinav Gupta, "Owner Incentives and Performance in Healthcare: Private Equity Investment in Nursing Homes," *The Review of Financial Studies*, 2023, p. hhad082.

Haushofer, Johannes and Jeremy Shapiro, "Household response to income changes: Evidence from an unconditional cash transfer program in Kenya," *Massachusetts Institute of Technology*, 2013, *24* (5), 1–57.

Henderson, J. Vernon, Adam Storeygard, and David N. Weil, "A bright idea for measuring economic growth," *American Economic Review*, 2011, *101* (4).

Higgins, Sean, "Financial Technology Adoption: Network Externalities of Cashless Payments in Mexico," *American Economic Review forthcoming*, 2020.

IIPS and ICF, "Demographic and Health Survey India, 2015-2016," International Institute for Population Sciences and ICF, https://dhsprogram.com/data/dataset/India_Standard-DHS_2015.cfm?flag=0, 2017. (Accessed March 12, 2019).

Kanz, Martin, "What does debt relief do for development? Evidence from India's bailout for rural households," *American Economic Journal: Applied Economics*, 2016, *8* (4), 66–99.

Karlan, Dean and Jonathan Zinman, "Expanding credit access: Using randomized supply decisions to estimate the impacts," *The Review of Financial Studies*, 2010, *23* (1), 433–464.

Kidane, Gebreyesus and Richard H Morrow, "Teaching mothers to provide home treatment of malaria in Tigray, Ethiopia: a randomised trial," *The lancet*, 2000, *356* (9229), 550–555.

Kremer, Michael and Rachel Glennerster, "Improving health in developing countries: evidence from randomized evaluations," in "Handbook of health economics," Vol. 2, Elsevier, 2011, pp. 201–315.

— **, Jessica Leino, Edward Miguel, and Alix Peterson Zwane**, "Spring cleaning: Rural water impacts, valuation, and property rights institutions," *The Quarterly Journal of Economics*, 2011, *126* (1), 145–205.

Limodio, Nicola, "Terrorism financing, recruitment, and attacks," *Econometrica*, 2022, *90* (4), 1711–1742.

Liu, Tong, "Bargaining with private equity: Implications for hospital prices and patient welfare," *Available at SSRN 3896410*, 2022.

Luby, Stephen P, Mubina Agboatwalla, Daniel R Feikin, John Painter, Ward Billhimer, Arshad Altaf, and Robert M Hoekstra, "Effect of handwashing on child health: a randomised controlled trial," *The Lancet*, 2005, *366* (9481), 225–233.

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

Moscoe, Ellen, Jacob Bor, and Till Bärnighausen, "Regression discontinuity designs are underutilized in medicine, epidemiology, and public health: a review of current and best practice," *Journal of Clinical Epidemiology*, 2015, *68* (2), 132–143.

ORGCC, "Population Census 2001," Office of the Registrar General Census Commissioner, India, https://censusindia.gov.in/census.website/data/census-tables, 2008. (Accessed January 11, 2019).

— , "Population Census 2011," Office of the Registrar General Census Commissioner, India, https://censusindia.gov.in/census.website/data/census-tables, 2014. (Accessed January 11, 2019).

Prina, Silvia, "Banking the poor via savings accounts: Evidence from a field experiment," *Journal of Development Economics*, 2015, *115*, 16–31.

Ramadorai, Tarun, "Report of the Household Finance Committee," Technical Report, Technical Report, Reserve Bank of India 2017.

RBI, "RBI Bank Branch Statistics, District-Wise Number of Functioning Offices of Commercial Banks," https://rbi.org.in/Scripts/AnnualPublications.aspx?head=Branch%20Banking%20Statistics, 2018. (Accessed June 19, 2018; note that the original data is not available anymore on the website. Instead, the RBI now links to their database on the Indian Economy. Data on this website is continuously updated, e.g., due to district border changes, and can thus differ from the original data. Original data can be made available upon request.).

_ , "RBI Master Office File," https://www.rbi.org.in/Scripts/query.aspx, 2018. (Accessed June 19, 2018; note that the original data is not available anymore on the website. Instead, the RBI developed a new query tool. Data accessible by this tool is continuously updated, e.g., due to district border changes, and can thus differ from the original data. Original data can be made available upon request.).

_ , "RBI Quarterly and Annual Statistics on Deposits and Credit of Scheduled Commercial Banks," https://rbi.org.in/Scripts/QuarterlyPublications.aspx?head=Quarterly%20Statistics%20on%20Deposits%20and%20Credit%20of%20Scheduled%20Commercial%20Banks, 2018. (Accessed June 19, 2018; note that the original data is not available anymore on the website. Instead, the RBI now links to their database on the Indian Economy. Data on this website is continuously updated, e.g., due to district border changes, and can thus differ from the original data. Original data can be made available upon request.).

Roux, Nicolás De and Nicola Limodio, "Deposit insurance and depositor behavior: Evidence from Colombia," *The Review of Financial Studies*, 2023, *36* (7), 2721–2755.

Sazawal, Sunil and Robert E Black, "Effect of pneumonia case management on mortality in neonates, infants, and preschool children: a meta-analysis of community-based trials," *The Lancet infectious diseases*, 2003, *3* (9), 547–556.

Stein, Luke CD and Constantine Yannelis, "Financial inclusion, human capital, and wealth accumulation: Evidence from the freedman's savings bank," *The Review of Financial Studies*, 2020, *33* (11), 5333–5377.

Tarozzi, Alessandro, Jaikishan Desai, and Kristin Johnson, "The impacts of microcredit: Evidence from Ethiopia," *American Economic Journal: Applied Economics*, 2015, *7* (1), 54–89.

Waddington, Hugh, Birte Snilstveit, Howard White, and Lorna Fewtrell, "Water, sanitation and hygiene interventions to combat childhood diarrhoea in developing countries," *New Delhi: International Initiative for Impact Evaluation*, 2009.

Young, Nathaniel, "Banking and growth: Evidence from a regression discontinuity analysis," 2017. EBRD Working Paper 207.

Appendix

Tables

Table A1: Branch Summary Statistics

	All districts				[-3,000;+3,000]			
	1997	2004	2010	2016	1997	2004	2010	2016
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Branch licenses (no.)	65	72	103	115	73	80	117	132
	(67)	(78)	(120)	(141)	(59)	(67)	(102)	(120)
Branches (no.)	65	71	103	171	73	79	116	198
	(68)	(76)	(116)	(185)	(59)	(66)	(100)	(166)
Observations	581	581	581	581	199	199	199	199

Standard deviations in parentheses. Data RBI. District level. All variables are winsorized at the 1st and 99th percentile. Regional rural banks are excluded.

Table A2: Households Summary Statistics (IHDS)

	IHDS I 2004/2005		IHDS II 2011/2012	
	All districts (1)	[-3,000,+3,000] (2)	All districts (3)	[-3,000,+3,000] (4)
Consumption				
Total consumption (Rs)	837	809	2,189	2,123
	(693)	(659)	(1,823)	(1,706)
Food consumption (Rs)	399	386	913	893
	(208)	(197)	(482)	(453)
Meals per day (no.)	2.83	2.86	2.75	2.78
	(0.67)	(0.67)	(0.58)	(0.57)
Financial Access				
Savings account (yes/no)			0.57	0.53
			(0.49)	(0.50)
Any loan (yes/no)	0.44	0.45	0.54	0.56
	(0.50)	(0.50)	(0.50)	(0.50)
Any bank loan (yes/no)			0.22	0.23
			(0.41)	(0.42)
Largest loan from bank (yes/no)	0.12	0.12	0.17	0.17
	(0.33)	(0.33)	(0.38)	(0.38)
Largest loan amt (Rs)	15,157	16,061	41,260	42,089
	(42,283)	(43,164)	(95,874)	(92,223)
Health				
Non-chronic: any illness (yes/no)	0.47	0.45	0.55	0.52
	(0.50)	(0.50)	(0.50)	(0.50)
Non-chronic: days ill (no.)	4.77	4.07	5.36	4.79
	(8.00)	(7.24)	(8.22)	(7.68)
Non-chronic: days missed (no.)	2.87	2.52	3.31	3.12
	(5.79)	(5.37)	(6.26)	(6.06)
Chronic: any illness (yes/no)	0.27	0.27	0.41	0.37
	(0.44)	(0.44)	(0.49)	(0.48)
Chronic: days missed (no.)	17.19	17.19	19.72	18.90
	(62.20)	(63.55)	(61.13)	(60.70)
Observations	39,584	16,184	41,703	16,965

Standard deviations in parentheses. Data IHDS I (2004/2005) and IHDS II (2011/2012). Household level. Variables in Rs or days are winsorized at the 1st and 99th percentile. No entry if not available in IHDS I. Amounts in Indian rupees are not inflation adjusted; inflation was 70 percent between 2004 and 2011. I observe that households in districts within the range of -3,000 to +3,000 of the normalized ratio are similar to households in all districts, strengthening external validity of my design.

Table A3: Households Summary Statistics (DHS)

	DHS 2015/2016	
	All districts (1)	[-3,000,+3,000] (2)
Health		
Non-chronic illness: sick child (yes/no)	0.27	0.26
	(0.45)	(0.44)
Health Care Visits		
Any illness: visit for sick child (yes/no)	0.11	0.10
	(0.31)	(0.30)
Any illness: visit for sick mother (yes/no)	0.16	0.16
	(0.37)	(0.37)
Generally go to: public provider (yes/no)	0.53	0.54
	(0.50)	(0.50)
Generally go to: private provider (yes/no)	0.44	0.44
	(0.50)	(0.50)
Generally go to: drug shop etc. (yes/no)	0.00	0.00
	(0.05)	(0.05)
Vaccinations		
Vaccinated child (yes/no)	0.85	0.86
	(0.36)	(0.35)
Healthcare Supply		
Big problem: distance to provider (yes/no)	0.36	0.34
	(0.48)	(0.47)
Big problem: transport to provider (yes/no)	0.34	0.32
	(0.47)	(0.47)
Big problem: no personnel (yes/no)	0.52	0.51
	(0.50)	(0.50)
Big problem: no female personnel (yes/no)	0.43	0.42
	(0.50)	(0.49)
Big problem: no drugs (yes/no)	0.53	0.52
	(0.50)	(0.50)
Observations	487,109	172,149

Standard deviations in parentheses. Data DHS (2015/2016). Household level.

Table A4: Economic Census District-Level Summary Statistics

	All districts		[-3,000,+3,000]	
	2005	2013	2005	2013
	(1)	(2)	(3)	(4)
Hospitals				
Hospitals (no.)	314	464	418	549
	(366)	(471)	(396)	(483)
Major source bank financing (yes/no)	0.02	0.02	0.03	0.01
	(0.03)	(0.02)	(0.03)	(0.02)
Other medical service providers				
Other medical service providers (no.)	448	546	494	556
	(658)	(829)	(628)	(772)
Major source bank financing (yes/no)	0.03	0.02	0.03	0.01
	(0.05)	(0.06)	(0.03)	(0.02)
All businesses				
All businesses (no.)	70,259	98,882	87,510	119,033
	(73,894)	(104,648)	(75,932)	(105,646)
Major source bank financing (yes/no)	0.03	0.02	0.03	0.02
	(0.03)	(0.02)	(0.02)	(0.01)
Observations	576	576	198	198

Standard deviations in parentheses. Data Economic Census. Household level. All variables in numbers are winsorized at the 1st and 99th percentile. Districts in the range of $\pm 3,000$ of the policy cutoff ratio have a slightly higher number of hospitals, other medical service providers, and all businesses.

Table A5: Economic Activity and Population Characteristics Are Smooth Pre-Policy

	1990	1991	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005
Nightlights														
Total light (log)			-0.52	-0.19	-0.21	-0.21	-0.17	-0.12	-0.13	-0.23	-0.20	-0.19	-0.12	-0.08
			(0.36)	(0.26)	(0.23)	(0.22)	(0.21)	(0.21)	(0.21)	(0.21)	(0.20)	(0.22)	(0.19)	(0.21)
Economic Census														
Empl. (log no.)	0.55						0.07							0.06
	(0.65)						(0.18)							(0.18)
Empl. manuf. (log no.)	0.20						-0.22							-0.09
	(0.56)						(0.24)							(0.22)
Empl. services (log no.)	0.60						0.26							0.10
	(0.66)						(0.19)							(0.16)
Population Census														
Pop. (log no.)		-0.13								0.02				
		(0.14)								(0.13)				
Pop. literate (log no.)		-0.17								-0.01				
		(0.15)								(0.13)				
Tar road (yes/no)		-43.87								28.74				
		(63.00)								(89.84)				
Observations	574	574	574	574	574	574	574	574	574	574	574	574	574	574

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are in parentheses and clustered on the district level. Bandwidth abbreviated by BW. Data SHRUG. District level. Combining different data sets, including night-light data, Economic Census data, and Population Census data. The unit of observation is town or village. I test whether units in treatment districts have, e.g., higher night light than units in control districts prior to the policy. The variables are defined as follows. Total light is the sum of the luminosity values of all pixels in a unit, obtained from the DMSP-OLS annual measures of nighttime luminosity. Employment measures the total employment, followed by a split by manufacturing and services. The Population Census measures total population, total literate population, and whether there is a tar road.

Table A6: Negligible Migration

	Migrated 5 years ago from other district (yes/no) (1)	Migrated anytime in past 90 years from other district (yes/no) (2)	Migrated 5 years ago from anywhere (yes/no) (3)
Treated	0.01	0.05	0.01
	(0.00)	(0.04)	(0.01)
Control Mean	0.00	0.11	0.01
First Stage	0.54	0.66	0.61
Bandwidth	1,633	2,363	1,982
Observations in BW	8,104	12,862	9,783
Total Observations	34,415	36,805	34,832
Baseline Control	Yes	Yes	Yes

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are in parentheses and clustered on the district level. Bandwidth abbreviated by BW. Data IHDS II (2011/2012). Household level.

Table A7: Other Policies Do Not Confound Results (1/2)

	NHM (yes/no) (1)	ICDS (1st wave) (yes/no) (2)	NREGA (1st wave) (yes/no) (3)	NREGA (2nd wave) (yes/no) (4)	RSBY (yes/no) (5)
Treated	0.22	-0.16	-0.23	-0.03	-0.04
	(0.24)	(0.15)	(0.20)	(0.25)	(0.25)
Control Mean	0.17	0.24	0.16	0.23	0.55
First Stage	0.69	0.76	0.71	0.67	0.71
Bandwidth	2,557	3,995	2,876	2,333	2,806
Observations in BW	171	253	191	156	187
Observations	581	581	581	581	581
Baseline Control	No	No	No	No	No

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are in parentheses and clustered on the district level. Bandwidth abbreviated by BW. Data Ministry of Health and Family Welfare, Ministry of Women and Child Development, Ministry of Rural Development, and Ministry of Labour and Employment. District level. Regressions do not include state-level fixed effects.

Table A8: Other Policies Do Not Confound Results (2/2)

	NHM (1)	ICDS (1st wave) (2)	NREGA (1st wave) (3)	NREGA (2nd wave) (4)	RSBY (5)
All districts					
Total priority districts (no.)	169	180	196	125	355
Total priority districts (%)	29	31	34	22	61
Priority districts above cutoff (no.)	135	142	170	85	217
Priority districts above cutoff (%)	36	38	45	23	58
Priority districts below cutoff (no.)	34	38	26	40	138
Priority districts below cutoff (%)	17	19	13	20	67
Corr priority district and 1[above]	0.20	0.20	0.33	0.04	-0.09
Within BW (-3,000;3,000)					
Total priority districts (no.)	44	55	53	41	102
Total priority districts (%)	29	31	34	22	61
Priority districts above cutoff (no.)	26	33	39	20	53
Priority districts above cutoff (%)	23	30	35	18	48
Priority districts below cutoff (no.)	18	22	14	21	49
Priority districts below cutoff (%)	20	25	16	24	56
Corr priority district and 1[above]	0.04	0.05	0.22	-0.07	-0.08

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are in parentheses and clustered on the district level. Bandwidth abbreviated by BW. Data Ministry of Health and Family Welfare, Ministry of Women and Child Development, Ministry of Rural Development, and Ministry of Labour and Employment. District level. Percent refers to the number of total districts within a given category; e.g., for priority districts above cutoff (%) within bandwidth, they constitute 26 percent of all districts above the cutoff within bandwidth.

Table A9: Private Banks React Stronger

	Post-Policy (2010)	
	Private branches (log no.) (1)	Public branches (log no.) (2)
Treated	0.47***	0.12**
	(0.17)	(0.05)
Control Mean	2.77	4.30
First Stage	0.80	0.80
Bandwidth	2,963	3,115
Obs. in BW	195	205
Observations	561	561
Baseline Control	Yes	Yes

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are in parentheses and clustered on the district level. Bandwidth abbreviated by BW. Data RBI. District level. All variables are winsorized at the 1st and 99th percentile.

Table A10: Branch Entry is Economically Meaningful

	Deposit accounts (log no.) (1)	Credit amount (log no.) (2)	Private bank Credit amount (log no.) (3)
Treated	0.10*	0.14*	0.43***
	(0.06)	(0.07)	(0.15)
Control Mean	13.81	3.14	1.37
First Stage	0.79	0.79	0.79
Bandwidth	2,542	2,348	2,248
Obs. in BW	170	157	147
Observations	561	553	553
Baseline Control	Yes	Yes	Yes

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are in parentheses and clustered on the district level. Bandwidth abbreviated by BW. Data RBI. District level. All variables are winsorized at the 1st and 99th percentile. Total credit amount in Column 2 does not include regional rural banks, which were excluded from the policy.

Table A11: Placebo Test: Regional Rural Banks Do Not React to the Policy

	Post-Policy (2010)	
	Branch Licenses (log no.) (1)	Branches (log no.) (2)
Treated	-0.54	-0.08
	(0.48)	(0.48)
Control Mean	1.51	1.09
First Stage	0.80	0.80
Bandwidth	2,812	2,959
Observations in BW	187	195
Total Observations	561	561
Baseline Control	Yes	Yes

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are in parentheses and clustered on the district level. Bandwidth abbreviated by BW. Data RBI. District level. All variables are winsorized at the 1st and 99th percentile. Only regional rural banks are analyzed.

Table A12: Health studies

Study	Context	Treatment	Duration	Outcome	Effect size
Waddington et al. (2009) (International Initiative for Impact Evaluation)	Meta-analysis	water, sanitation, hygiene interventions	5 months - 2 years	child diarrhea in past weeks	31-42% decrease
Gertler (2004) (AER)	Mexico	conditional cash program	2 years	child non-chronic diseases past month	27% decrease
Kremer et al. (2011) (QJE)	Kenya	spring water protection	3 years	child diarrhea past week	25% decrease
Luby et al. (2005) (Lancet)	Pakistan	hand washing promotion	1 year	child diarrhea	53% decrease
Sazawal and Black (2003) (Lancet infectious diseases)	Meta-analysis	pneumonia case management	various	child mortality	24% decrease
Björkman and Svensson (2009) (QJE)	Uganda	community-based monitoring of healthcare providers	1 year	child mortality	33% decrease
Björkman-Nykvist et al. (2014) (working paper)	Uganda	community health workers	3 years	child mortality	27% decrease
Kidane and Morrow (2000) (Lancet)	Ethiopia	education for mothers to detect and treat malaria	1 year	child mortality	40% decrease
Bang et al. (1999) (Lancet)	India	home-based neonatal care	2 years	child mortality	46% decrease

Table A13: Smooth Health Status Pre-Policy

	Non-chronic illness			Chronic illness	
	Any (yes/no) (1)	Days ill (log no.) (2)	Days missed (log no.) (3)	Any (yes/no) (4)	Days missed (log no.) (5)
Treated	-0.07	-0.21	-0.28	0.03	-0.07
	(0.06)	(0.17)	(0.19)	(0.05)	(0.19)
Control Mean	0.40	0.83	0.63	0.26	0.70
First Stage	0.71	0.69	0.66	0.65	0.66
Bandwidth	4,363	3,812	2,625	2,482	2,659
Obs. in BW	18,207	16,061	11,553	11,296	12,006
Observations	31,913	31,913	31,794	31,794	31,794
Baseline Control	No	No	No	No	No

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are in parentheses and clustered on the district level. Bandwidth abbreviated by BW. Data IHDS I (2004/2005). Household level. All variables measured in days are in log and winsorized at the 1st and 99th percentile. Non-chronic illnesses include fever, diarrhea, and cough in the past 30 days. Chronic illnesses include, for instance, heart disease and cancer, ever diagnosed (column 4) or days unable to work in the past 12 months (column 5).

Table A14: Results Hold With Baseline Control

	Non-chronic illness			Chronic illness	
	Any (yes/no) (1)	Days ill (log no.) (2)	Days missed (log no.) (3)	Any (yes/no) (4)	Days missed (log no.) (5)
Treated	-0.20**	-0.44**	-0.57***	-0.01	-0.28
	(0.08)	(0.20)	(0.17)	(0.06)	(0.18)
Control Mean	0.53	1.10	0.77	0.39	1.04
First Stage	0.63	0.64	0.67	0.60	0.63
Bandwidth	2,327	2,376	2,922	2,045	2,256
Obs. in BW	12,967	13,099	16,453	10,160	12,544
Observations	31,710	31,710	31,794	30,179	31,710
Baseline Control	Yes	Yes	Yes	Yes	Yes

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are in parentheses and clustered on the district level. Bandwidth abbreviated by BW. Data IHDS II (2004/2005). Household level. All variables measured in days are in log and winsorized at the 1st and 99th percentile. Non-chronic illnesses include fever, diarrhea, and cough in the past 30 days. Chronic illnesses include, for instance, heart disease and cancer, ever diagnosed (column 4) or days unable to work in the past 12 months (column 5).

Table A15: Results Robust to Different Transformations

	Non-chronic illness (days ill)					
	log(1+x) winsorized (1)	log(1+x) not winsorized (2)	ihs winsorized (3)	ihs not winsorized (4)	level winsorized (5)	level not winsorized (6)
Treated	-0.43**	-0.43**	-0.53**	-0.53**	-2.68*	-2.80*
	(0.19)	(0.19)	(0.23)	(0.24)	(1.42)	(1.56)
Control Mean	1.08	1.09	1.35	1.36	4.88	5.00
First Stage	0.66	0.66	0.66	0.66	0.64	0.63
Bandwidth	2,312	2,303	2,314	2,306	2,139	2,079
Obs. in BW	12,927	12,834	12,927	12,927	11,222	10,518
Observations	38,375	38,375	38,375	38,375	36,673	36,673
Baseline Control	No	No	No	No	No	No

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are in parentheses and clustered on the district level. Bandwidth abbreviated by BW. Data IHDS II (2011/2012). Household level. Since the number of days that household members are ill is zero for 45% of households, I show robustness for different transformations. The winsorization is at the 1st and 99th percentile. IHS is the inverse hyperbolic sine transformation (ihs(x) = ln(x + $\sqrt{(x^2+1)}$)). Results are robust to different transformations.

Table A16: No Evidence of Bias Due to Self-Reporting

	Vaccinated child (yes/no) (1)
Treated	0.07*
	(0.04)
Control Mean	0.86
First Stage	0.72
Bandwidth	2,898
Obs. in BW	26,117
Observations	86,079
Baseline Control	No

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are in parentheses and clustered on the district level. Bandwidth abbreviated by BW. Data DHS (2015/2016). Household level.

Table A17: Financial Access is Smooth Pre-Policy

	Any loan (yes/no) (1)	Largest loan amount (log Rs) (2)	Largest loan from bank (yes/no) (3)
Treated	0.02	0.39	-0.02
	(0.09)	(0.78)	(0.02)
Control Mean	0.40	3.92	0.12
First Stage	0.69	0.70	0.71
Bandwidth	3,821	3,862	4,325
Obs. in BW	16,183	16,395	18,090
Observations	31,911	31,913	31,912
Baseline Control	No	No	No

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are in parentheses and clustered on the district level. Bandwidth abbreviated by BW. Data IHDS I (2004/2005). Household level. Variable in Rs is transformed to log and winsorized at the 1st and 99th percentile.

Table A18: Business Activity is Smooth Pre-Policy

	Total employment		
	Total (log nr.) (1)	Manufacturing (log nr.) (2)	Services (log nr.) (3)
Treated	0.01	-0.12	0.04
	(0.12)	(0.20)	(0.11)
Control Mean	11.57	10.09	11.23
First Stage	0.78	0.77	0.79
Bandwidth	2,713	2,635	2,985
Obs. in BW	180	171	195
Observations	555	555	555
Baseline Control	No	No	No

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are in parentheses and clustered on the district level. Bandwidth abbreviated by BW. Data Economic Census (2005). District level. Variables in log and winsorized at the 1st and 99th percentile.

Table A19: No Effect on State Expenditure

	Medical and Public Health (log lakh Rs) (1)	Water supply and Sanitation (log lakh Rs) (2)	Nutrition (log lakh Rs) (3)
Treated	-0.14	0.06	-0.43
	(0.25)	(0.30)	(0.60)
Control Mean	11.95	10.58	10.75
First Stage	0.71	0.74	0.76
Bandwidth	2,769	3,397	3,377
Obs. in BW	182	213	194
Observations	570	570	470
Baseline Control	No	No	No

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are in parentheses and clustered on the district level. Bandwidth abbreviated by BW. Data RBI (2010). Variable in lakh (= hundred thousand) Rs and transformed to log plus winsorized at the 1st and 99th percentile.

Table A20: Healthcare Activity is Smooth Pre-Policy

	Healthcare Providers	
	Institutional loan (share)	Number (log nr.)
	(1)	(2)
Treated	0.00	-0.15
	(0.01)	(0.16)
Control Mean	0.03	5.42
First Stage	0.79	0.80
Bandwidth	2,638	4,328
Obs. in BW	173	273
Observations	556	557
Baseline Control	No	No

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are in parentheses and clustered on the district level. Bandwidth abbreviated by BW. Data Economic Census (2005). District level. Variables in log and winsorized at the 1st and 99th percentile.

Table A21: Shift Towards Private Providers

	Generally go for treatment to		
	Government provider (yes/no)	Private provider (yes/no)	Shop or stay home (yes/no)
	(1)	(2)	(3)
Treated	-0.06**	0.10***	-0.00
	(0.03)	(0.03)	(0.00)
Control Mean	0.52	0.45	0.00
First Stage	0.73	0.71	0.69
Bandwidth	2,898	2,648	2,262
Observations in BW	202,459	184,429	156,853
Total Observations	577,928	577,928	566,715
Baseline Control	No	No	No

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are in parentheses and clustered on the district level. Bandwidth abbreviated by BW. Data DHS (2015/2016). Household level.

Table A22: Households Spend More on Food and Hygiene

	Low fixed cost		High fixed cost	
	Hygiene expenses (log Rs) (1)	Meals per day (no.) (2)	Fridge (yes/no) (3)	Toilet (rank) (4)
Treated	0.24**	0.17**	-0.01	-0.08
	(0.12)	(0.08)	(0.07)	(0.14)
Control Mean	2.74	4.03	0.21	2.15
First Stage	0.65	0.65	0.58	0.70
Bandwidth	2,266	2,193	1,837	3,426
Obs. in BW	12,458	11,974	9,563	18,088
Observations	38,045	36,640	36,432	38,397
Baseline Control	No	No	No	No

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are in parentheses and clustered on the district level. Bandwidth abbreviated by BW. Data IHDS (2011/2012). Household level. Hygiene expenses in the past month in log rupees and winsorized at the 1st and 99th percentile. The toilet variable is a rank where 1 is no toilet, 2 is the traditional pit latrine, 3 is a semi-flush latrine, and 4 is a flush toilet.

Table A23: Suggestive Evidence of Higher Healthcare Demand

	Vaccinated child (yes/no) (1)	Birth in healthcare facility (yes/no) (2)
Treated	0.072*	0.005***
	(0.040)	(0.002)
Control Mean	0.86	0.02
First Stage	0.72	0.72
Bandwidth	2,898	3,023
Obs. in BW	26,117	172,892
Observations	86,079	471,985
Baseline Control	No	No

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are in parentheses and clustered on the district level. Bandwidth abbreviated by BW. Data DHS (2015/2016). Household level. These indicators for healthcare utilization are indirect measures of healthcare demand that – unlikely medical expenditure or visits for diseases – are not likely to decrease with improved health status.

Table A24: Robustness to Different Bandwidth Multipliers

	Bandwidth Multiplier						
	x0.50	**x0.75**	**x1.00**	**x1.25**	**x1.50**	**x1.75**	**x2.00**
Banks *(Table 2)*							
Branch licenses 2010 (log no.)	0.18**	0.23***	0.19***	0.17***	0.15***	0.13***	0.13***
	(0.07)	(0.06)	(0.05)	(0.05)	(0.05)	(0.05)	(0.04)
Branches 2010 (log no.)	0.19**	0.20***	0.17***	0.14**	0.11**	0.11**	0.12***
	(0.08)	(0.06)	(0.06)	(0.05)	(0.05)	(0.05)	(0.05)
Household health *(Tables 3 and 4)*							
Non-chronic: any illness (yes/no)	-0.48	-0.26	-0.19**	-0.18***	-0.16**	-0.15**	-0.13**
	(0.34)	(0.16)	(0.08)	(0.07)	(0.06)	(0.06)	(0.06)
Non-chronic: days ill (log no.)	-1.12	-0.57	-0.43**	-0.42***	-0.39***	-0.34**	-0.29**
	(0.83)	(0.36)	(0.19)	(0.16)	(0.15)	(0.14)	(0.13)
Non-chronic: days missed (log no.)	-1.27	-0.78**	-0.61***	-0.56***	-0.51***	-0.43***	-0.36***
	(0.82)	(0.34)	(0.20)	(0.16)	(0.15)	(0.14)	(0.13)
Chronic: any illness (yes/no)	-0.28	-0.01	-0.00	-0.03	-0.03	-0.03	-0.04
	(0.23)	(0.09)	(0.05)	(0.04)	(0.04)	(0.04)	(0.04)
Chronic: days missed (log no.)	-1.16	-0.55	-0.31	-0.36**	-0.39***	-0.37***	-0.36***
	(0.80)	(0.38)	(0.20)	(0.15)	(0.15)	(0.14)	(0.13)
Non-chronic: Sick child (yes/no)	-0.12***	-0.10***	-0.06*	-0.04	-0.03	-0.02	-0.02
	(0.04)	(0.04)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Any illness: visit for sick child (yes/no)	-0.07***	-0.04**	-0.02*	-0.02	-0.01	-0.01	-0.01
	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Any illness: visit for sick mother (yes/no)	-0.11**	-0.07**	-0.05*	-0.04	-0.02	-0.01	-0.00
	(0.05)	(0.03)	(0.03)	(0.02)	(0.02)	(0.02)	(0.02)
Banking activity *(Tables 5, 6, and 7)*							
Households: savings account (yes/no)	0.21	0.24*	0.19**	0.18**	0.13*	0.09	0.06
	(0.22)	(0.13)	(0.10)	(0.08)	(0.07)	(0.06)	(0.06)
Households: bank loan (yes/no)	-0.19	-0.04	0.04	0.05	0.04	0.03	0.02
	(0.15)	(0.07)	(0.05)	(0.04)	(0.04)	(0.04)	(0.04)
Employment (log no.)	0.08	0.15**	0.12*	0.08	0.06	0.05	0.05
	(0.09)	(0.07)	(0.07)	(0.06)	(0.06)	(0.06)	(0.06)
Hospitals (log no.)	1.74**	1.27***	0.89***	0.62**	0.42*	0.28	0.19
	(0.76)	(0.45)	(0.33)	(0.27)	(0.24)	(0.22)	(0.19)

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are in parentheses and clustered on the district level. Bandwidth abbreviated by BW. For details of the regression, refer to the respective main table. Summarized in Figure A6.

Table A25: Robustness to Different Bandwidth Selectors (1/2)

	MSE-optimal		CER-optimal	
	Common	**Two-sided**	**Common**	**Two-sided**
Banks *(Table 2)*				
Branch licenses 2010 (log no.)	0.19***	0.20***	0.23***	0.18***
	(0.05)	(0.06)	(0.07)	(0.06)
	0.24***	0.27***	0.26***	0.22***
	(0.05)	(0.06)	(0.07)	(0.06)
	0.24***	0.27***	0.26***	0.22***
	(0.06)	(0.07)	(0.07)	(0.07)
Branches 2010 (log no.)	0.17***	0.17***	0.20***	0.17**
	(0.06)	(0.06)	(0.07)	(0.07)
	0.21***	0.24***	0.22***	0.20***
	(0.06)	(0.06)	(0.07)	(0.07)
	0.21***	0.24***	0.22***	0.20***
	(0.07)	(0.07)	(0.07)	(0.07)
Household health *(Table 3)*				
Non-chronic: any illness (yes/no)	-0.19**	-0.13*	-0.26	-0.17
	(0.08)	(0.08)	(0.17)	(0.15)
	-0.21**	-0.16**	-0.28*	-0.19
	(0.08)	(0.08)	(0.17)	(0.15)
	-0.21**	-0.16	-0.28	-0.19
	(0.10)	(0.10)	(0.19)	(0.16)
Non-chronic: days ill (log no.)	-0.43**	-0.36*	-0.58	-0.65
	(0.19)	(0.20)	(0.36)	(0.43)
	-0.48**	-0.43**	-0.61*	-0.71*
	(0.19)	(0.20)	(0.36)	(0.43)
	-0.48**	-0.43*	-0.61	-0.71
	(0.23)	(0.26)	(0.41)	(0.50)
Non-chronic: days missed (log no.)	-0.61***	-0.56***	-0.79**	-0.88*
	(0.20)	(0.22)	(0.34)	(0.47)
	-0.73***	-0.68***	-0.87**	-0.97**
	(0.20)	(0.22)	(0.34)	(0.47)
	-0.73***	-0.68**	-0.87**	-0.97*
	(0.23)	(0.28)	(0.38)	(0.53)
Chronic: any illness (yes/no)	-0.00	-0.02	-0.01	-0.13
	(0.05)	(0.11)	(0.09)	(0.18)
	0.01	-0.02	-0.01	-0.14
	(0.05)	(0.11)	(0.09)	(0.18)
	0.01	-0.02	-0.01	-0.14
	(0.06)	(0.14)	(0.10)	(0.23)
Chronic: days missed (log no.)	-0.31	-0.72*	-0.55	-1.17
	(0.20)	(0.44)	(0.38)	(0.84)
	-0.35*	-0.83*	-0.59	-1.29
	(0.20)	(0.44)	(0.38)	(0.84)
	-0.35	-0.83	-0.59	-1.29
	(0.23)	(0.55)	(0.41)	(1.09)

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are in parentheses and clustered on the district level. Bandwidth abbreviated by BW. The first and second columns are MSE-optimal bandwidths, initially identical and then different to the left and right of the cutoff. The third and fourth columns indicate CER (coverage error rate)-optimal bandwidths, first identical and then different to the left and right of the cutoff (Calonico et al., 2020). In each parcel, I first report the conventional RD estimator with conventional variance estimator. Below is the bias-corrected RD estimator with the conventional variance estimator, followed by the bias-corrected RD estimator with robust variance estimator (Calonico et al., 2014). For details of the regression, refer to the respective main table. Summarized in Figures A8 and A12.

Table A26: **Robustness to Different Bandwidth Selectors (2/2)**

	MSE-optimal		CER-optimal	
	Common	**Two-sided**	**Common**	**Two-sided**
***Household health** (Table 4)*				
Non-chronic: Sick child (yes/no)	-0.06*	-0.04	-0.11***	-0.08*
	(0.03)	(0.03)	(0.04)	(0.04)
	-0.08**	-0.06*	-0.12***	-0.09**
	(0.03)	(0.03)	(0.04)	(0.04)
	-0.08*	-0.06	-0.12***	-0.09*
	(0.04)	(0.04)	(0.04)	(0.05)
Any illness: visit for sick child (yes/no)	-0.02*	-0.03*	-0.04**	-0.04**
	(0.01)	(0.01)	(0.02)	(0.02)
	-0.04**	-0.04***	-0.05***	-0.05***
	(0.01)	(0.01)	(0.02)	(0.02)
	-0.04*	-0.04**	-0.05**	-0.05***
	(0.02)	(0.02)	(0.02)	(0.02)
Any illness: visit for sick mother (yes/no)	-0.05*	-0.03	-0.08**	-0.06**
	(0.03)	(0.02)	(0.03)	(0.03)
	-0.08***	-0.05**	-0.09***	-0.08***
	(0.03)	(0.02)	(0.03)	(0.03)
	-0.08**	-0.05*	-0.09**	-0.08**
	(0.03)	(0.03)	(0.04)	(0.03)
***Banking activity** (Table 5, 6, and 7)*				
Households: savings account (yes/no)	0.19**	0.21*	0.24*	0.26
	(0.10)	(0.11)	(0.13)	(0.18)
	0.26***	0.29***	0.29**	0.31*
	(0.10)	(0.11)	(0.13)	(0.18)
	0.26**	0.29**	0.29**	0.31
	(0.12)	(0.14)	(0.14)	(0.20)
Households: bank loan (yes/no)	0.04	0.04	-0.05	-0.00
	(0.05)	(0.05)	(0.07)	(0.08)
	0.07	0.04	-0.04	-0.01
	(0.05)	(0.05)	(0.07)	(0.08)
	0.07	0.04	-0.04	-0.01
	(0.06)	(0.07)	(0.08)	(0.10)
Employment (log no.)	0.12*	0.15**	0.15**	0.14*
	(0.07)	(0.07)	(0.07)	(0.08)
	0.15**	0.21***	0.17**	0.18**
	(0.07)	(0.07)	(0.07)	(0.08)
	0.15*	0.21**	0.17**	0.18**
	(0.08)	(0.08)	(0.08)	(0.08)
Hospitals (log no.)	0.89***	0.79**	1.32***	1.18**
	(0.33)	(0.33)	(0.46)	(0.49)
	1.16***	1.06***	1.51***	1.37***
	(0.33)	(0.33)	(0.46)	(0.49)
	1.16***	1.06***	1.51***	1.37**
	(0.40)	(0.40)	(0.51)	(0.54)

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are in parentheses and clustered on the district level. Bandwidth abbreviated by BW. The first and second columns are MSE-optimal bandwidths, initially identical and then different to the left and right of the cutoff. The third and fourth columns indicate CER (coverage error rate)-optimal bandwidths, first identical and then different to the left and right of the cutoff (Calonico et al., 2020). In each parcel, I first report the conventional RD estimator with conventional variance estimator. Below is the bias-corrected RD estimator with the conventional variance estimator, followed by the bias-corrected RD estimator with robust variance estimator (Calonico et al., 2014). For details of the regression, refer to the respective main table. Summarized in Figures A8 and A12.

Table A27: Robustness to Different Polynomial Degrees

	Polynomial Degree		
	One	**Two**	**Three**
Banks *(Table 2)*			
Branch licenses 2010 (log no.)	0.19***	0.33***	0.46***
	(0.05)	(0.09)	(0.14)
Branches 2010 (log no.)	0.17***	0.31***	0.44***
	(0.06)	(0.09)	(0.14)
Household health *(Tables 3 and 4)*			
Non-chronic: any illness (yes/no)	-0.19**	-0.22*	-0.22
	(0.08)	(0.13)	(0.16)
Non-chronic: days ill (log no.)	-0.43**	-0.52*	-0.58
	(0.19)	(0.32)	(0.41)
Non-chronic: days missed (log no.)	-0.61***	-0.84**	-0.97*
	(0.20)	(0.38)	(0.52)
Chronic: any illness (yes/no)	-0.00	0.02	0.06
	(0.05)	(0.06)	(0.10)
Chronic: days missed (log no.)	-0.31	-0.41*	-0.38
	(0.20)	(0.23)	(0.32)
Non-chronic: Sick child (yes/no)	-0.06*	-0.08	-0.23**
	(0.03)	(0.05)	(0.11)
Any illness: visit for sick child (yes/no)	-0.02*	-0.06**	-0.13**
	(0.01)	(0.03)	(0.05)
Any illness: visit for sick mother (yes/no)	-0.05*	-0.11**	-0.17*
	(0.03)	(0.05)	(0.10)
Banking activity *(Tables 5, 6, and 7)*			
Households: savings account (yes/no)	0.19**	0.28*	0.29
	(0.10)	(0.16)	(0.21)
Households: bank loan (yes/no)	0.04	0.08	0.09
	(0.05)	(0.06)	(0.08)
Employment (log no.)	0.12*	0.18	0.22
	(0.07)	(0.11)	(0.17)
Hospitals (log no.)	0.89***	1.24**	1.77
	(0.33)	(0.57)	(1.26)

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are in parentheses and clustered on the district level. Bandwidth abbreviated by BW. For details of the regression, refer to the respective main table. Summarized in Figure A9.

Table A28: Placebo Cutoffs

	Placebo Cutoffs						
	-3,000	**-2,000**	**-1,000**	**0**	**1,000**	**2,000**	**3,000**
Banks *(Table 2)*							
Branch licenses 2010 (log no.)	0.92	0.01	0.22	0.00	0.78	0.06	0.04
Branches 2010 (log no.)	0.87	0.52	0.40	0.00	0.50	0.14	0.04
Household health *(Tables 3 and 4)*							
Non-chronic: any illness (yes/no)	0.79	0.94	0.34	0.02	0.26	0.03	0.55
Non-chronic: days ill (log no.)	0.82	0.82	0.27	0.02	0.17	0.01	0.75
Non-chronic: days missed (log no.)	0.88	0.29	0.23	0.00	0.23	0.03	0.84
Chronic: any illness (yes/no)	0.95	0.07	0.09	0.94	0.94	0.02	0.22
Chronic: days missed (log no.)	0.88	0.19	0.66	0.11	0.33	0.65	0.14
Non-chronic: Sick child (yes/no)	0.21	0.64	0.96	0.06	0.15	0.70	0.42
Any illness: visit for sick child (yes/no)	0.44	0.84	0.56	0.10	0.35	0.99	0.44
Any illness: visit for sick mother (yes/no)	0.46	0.34	0.84	0.06	0.51	0.65	0.55
Banking activity *(Tables 5, 6, and 7)*							
Households: savings account (yes/no)	0.91	0.69	0.12	0.05	0.32	0.32	0.85
Households: bank loan (yes/no)	0.99	0.60	0.36	0.37	0.81	0.21	0.18
Employment (log no.)	0.19	0.54	0.96	0.09	0.54	0.44	.
Hospitals (log no.)	0.59	0.97	0.01	0.01	0.16	0.81	0.96

P-values for different (placebo) cutoffs shown. For details of the regressions, refer to the respective main table. Summarized in Figure A10.

Table A29: Standard Error Adjustments

	Standard Error Adjustments			
	None	Multiple hypothesis testing	Spatial correlation (500km)	Spatial correlation (100km)
	(1)	(2)	(3)	(4)
Banks *(Table 2)*				
Branch licenses 2010 (log no.)	0.00	0.00	0.00	0.00
Branches 2010 (log no.)	0.00	0.01	0.00	.
Household health *(Tables 3 and 4)*				
Non-chronic: any illness (yes/no)	0.02	0.04	.	.
Non-chronic: days ill (log no.)	0.02	0.04	.	.
Non-chronic: days missed (log no.)	0.00	0.01	.	.
Chronic: any illness (yes/no)	0.94	0.16	.	.
Chronic: days missed (log no.)	0.11	0.08	.	.
Non-chronic: Sick child (yes/no)	0.06	0.06	.	.
Any illness: visit for sick child (yes/no)	0.10	0.08	.	.
Any illness: visit for sick mother (yes/no)	0.06	0.06	.	.
Banking activity *(Tables 5, 6, and 7)*				
Households: savings account (yes/no)	0.05	0.06	.	.
Households: bank loan (yes/no)	0.37	0.14	.	.
Employment (log no.)	0.09	0.08	0.23	0.01
Hospitals (log no.)	0.01	0.02	0.06	0.26

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are in parentheses and clustered on the district level. Bandwidth abbreviated by BW. Column 2 shows adjustments to multiple hypothesis testing (false discovery rate), Columns 3 and 4 to spatial correlation (Conley standard errors, 100km, and 500km). For details of the regression, refer to the respective main table. Summarized in Figure A11.

Figures

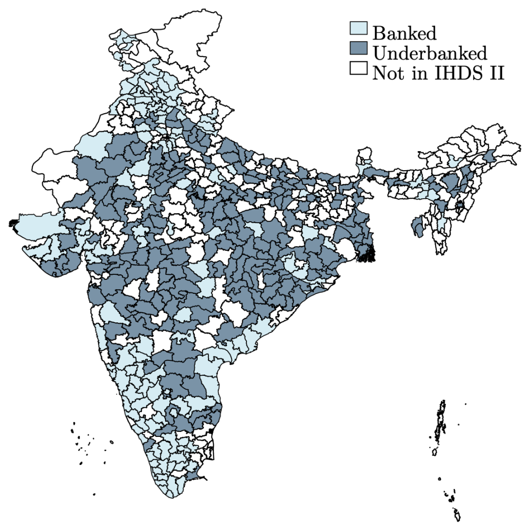


Figure A1. Districts Interviewed. In IHDS II, interviews were conducted in 65 percent of all districts.

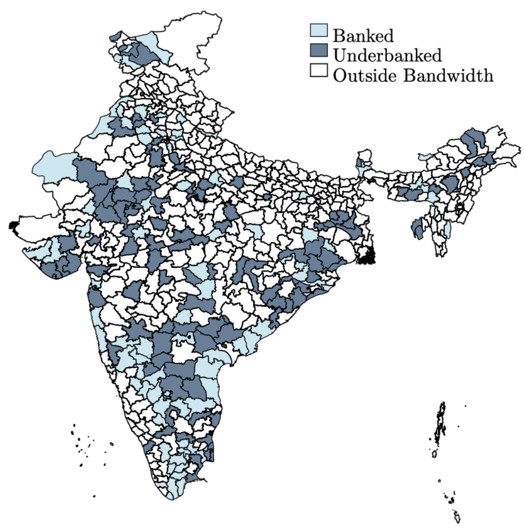


Figure A2. Districts With a Population-to-Branch Ratio Within Typical Bandwidth. There are 111 districts underbanked and 88 districts banked within the typical bandwidth of $\pm$ 3,000.

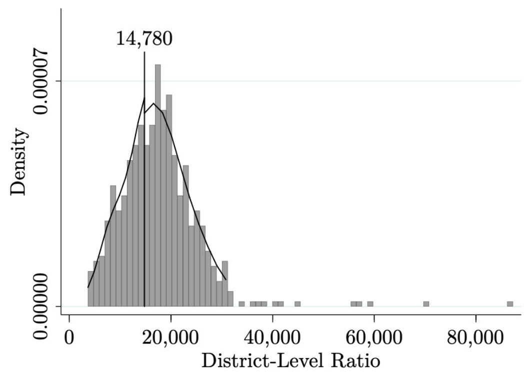


Figure A3. McCrary (2008) Density Test. There is no evidence of manipulation around the cutoff. The McCrary estimator is -0.1998 with a p-value of 0.8416; I do not reject smoothness around the cutoff.

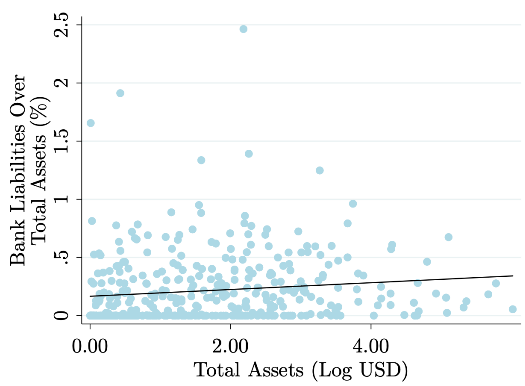
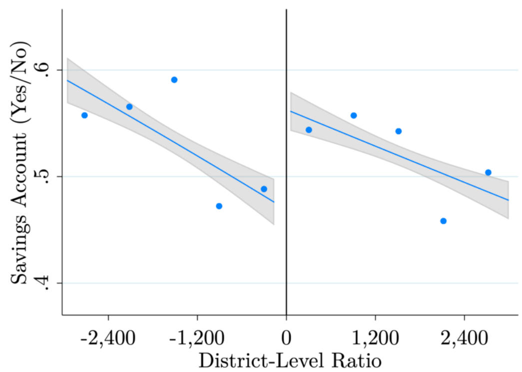
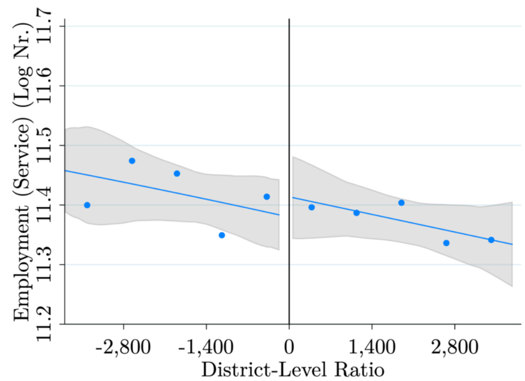


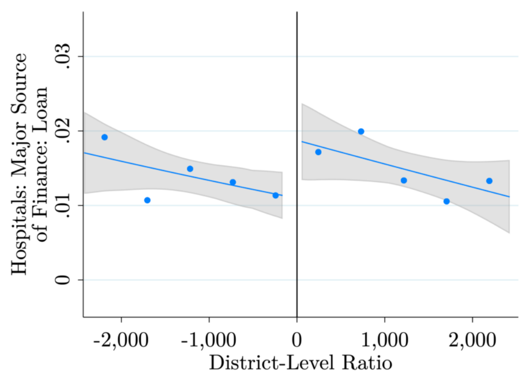
Figure A4. Relationship Between Bank Liabilities and Total Assets. As expected, there is a positive relationship between the share of bank liabilities over total assets and the size of the company proxied by total assets. However, there are many companies of lower asset size that have a relatively high share of bank liabilities over total assets.



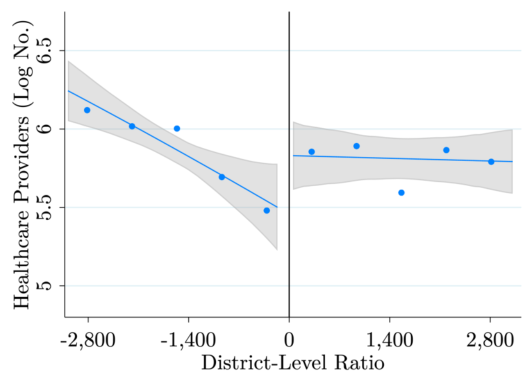
(a) Post-Policy (2011/2012)



(b) Post-Policy (2011/2012)

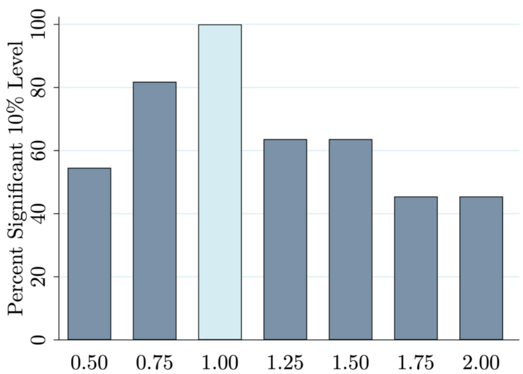


(c) Post-Policy (2011/2012)

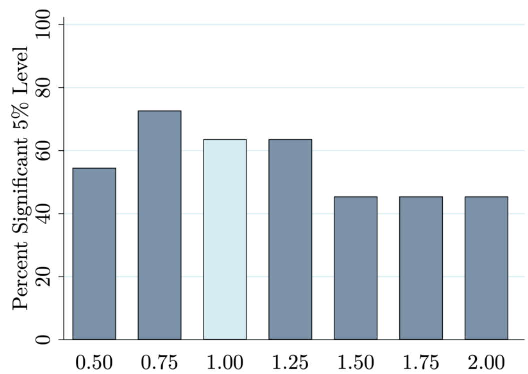


(d) Post-Policy (2011/2012)

Figure A5. Mechanisms. These graphs show binned means to the left and right of the cutoff within the optimal bandwidth. They also show local linear polynomials to the left and right of the cutoff, with 95 percent confidence intervals in gray.

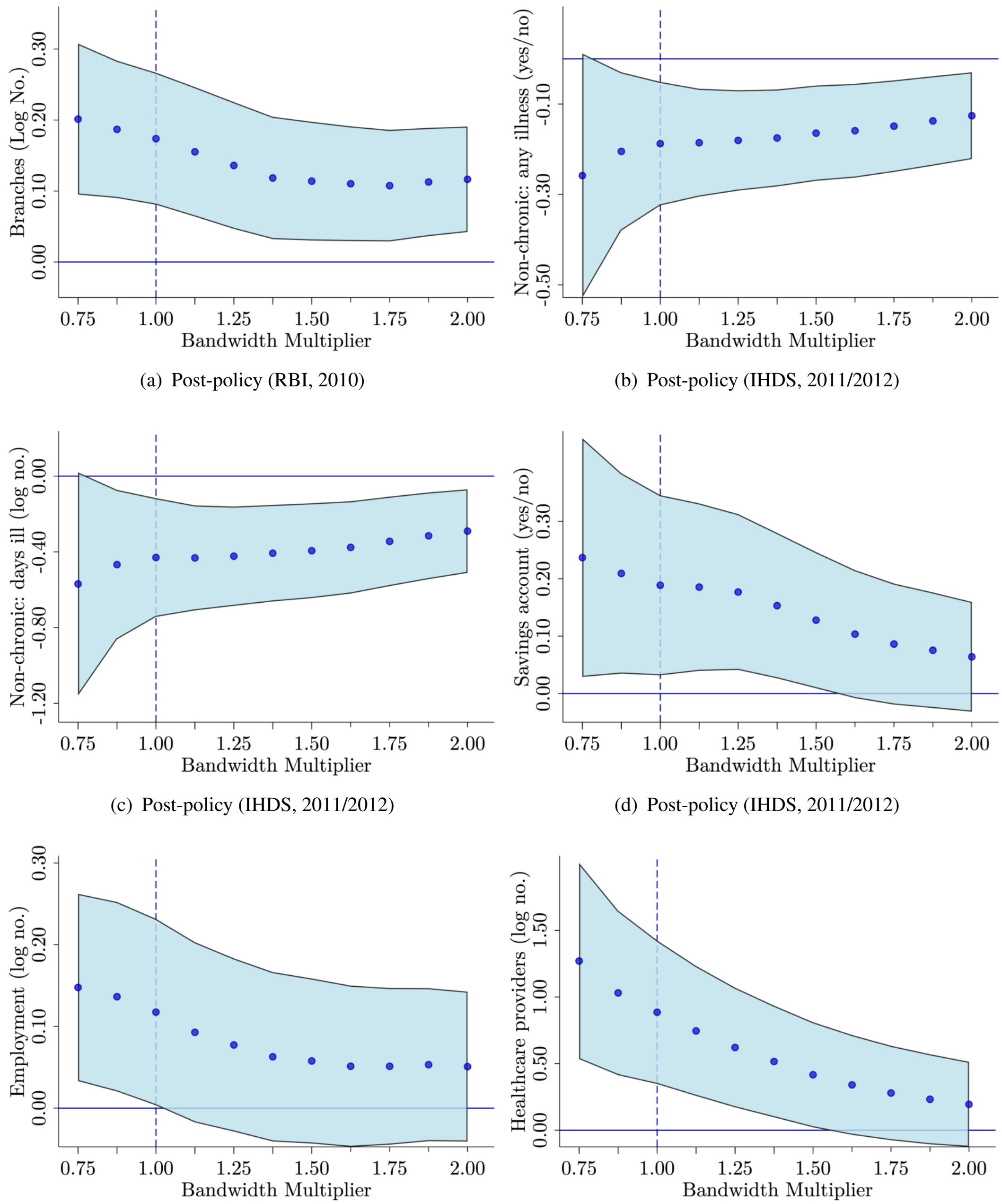


(a) 10% Significance



(b) 5% Significance

Figure A6. Percent of Results That Remain Significant Under Different Bandwidth Multipliers. Light blue indicates the main specification (optimal bandwidth), dark blue indicates alternative specifications (optimal bandwidth multiplied by factor, e.g., 1.25 times optimal bandwidth). Refers to Table A24.



(a) Post-policy (RBI, 2010)

(b) Post-policy (IHDS, 2011/2012)

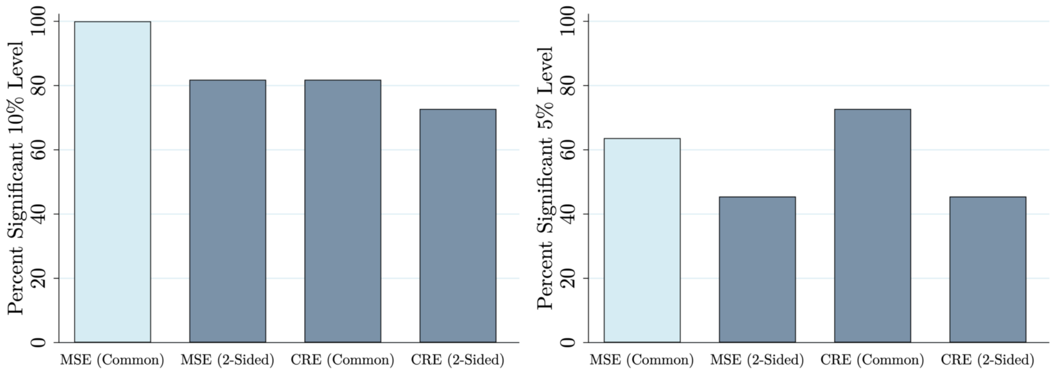
(c) Post-policy (IHDS, 2011/2012)

(d) Post-policy (IHDS, 2011/2012)

(e) Post-policy (EC, 2013)

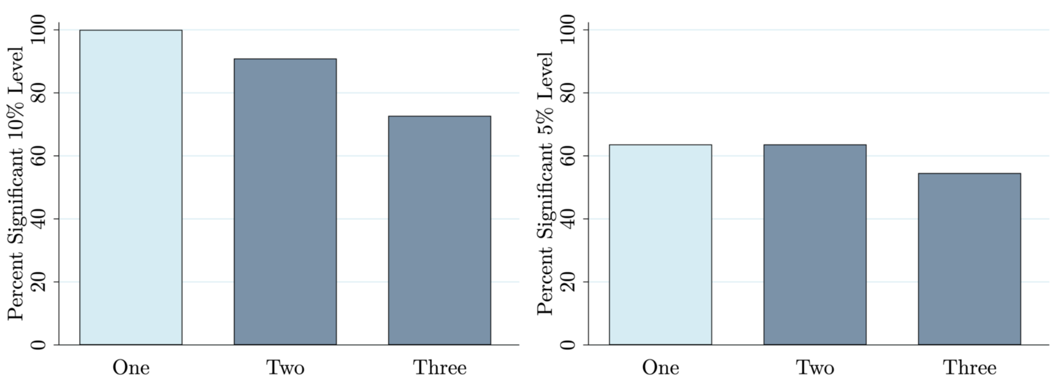
(f) Post-policy (EC, 2013)

Figure A7. Robustness Under Different Bandwidth Multipliers. 90 percent confidence intervals. Refers to Table A24.



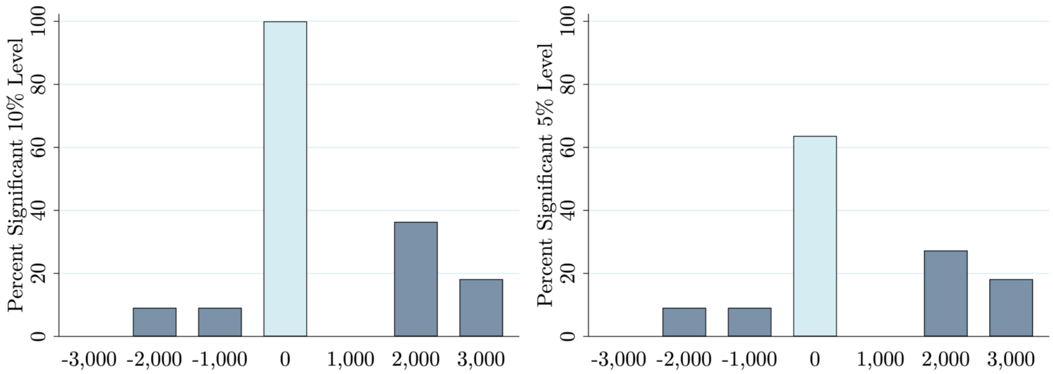
(a) 10% Significance (b) 5% Significance

Figure A8. Percent of Results That Remain Significant Under Different Bandwidth Selectors. Light blue indicates the main bandwidth (MSE-optimal with common bandwidth to the left and to the right of the cutoff), and dark blue indicates alternative bandwidths. The second column indicates MSE-optimal bandwidths different to the left and to the right of the cutoff. This is followed by coverage error rate (CER)-optimal bandwidths, first common bandwidth and then different to the left and right of the cutoff (Calonico et al., 2020). Refers to Tables A25 and A26.



(a) 10% Significance (b) 5% Significance

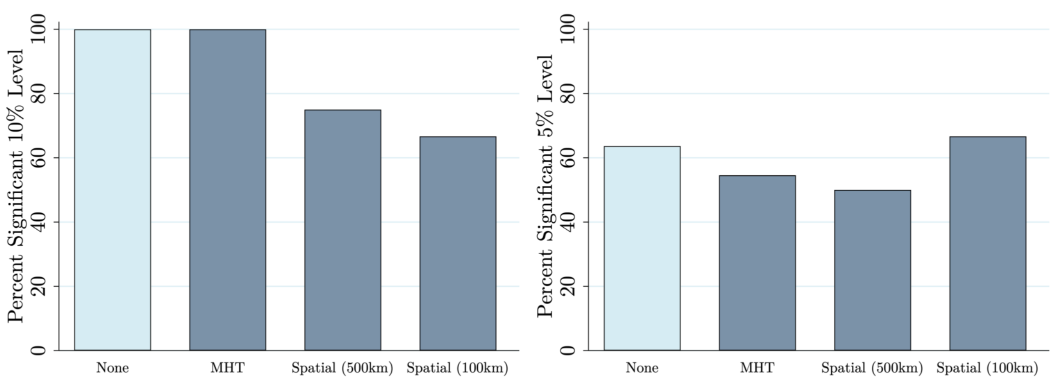
Figure A9. Percent of Results That Remain Significant Under Different Polynomial Degrees. Light blue indicates the main specification (degree one), dark blue indicates alternative specifications (degree two and three). Refers to Table A27.



(a) 10% Significance

(b) 5% Significance

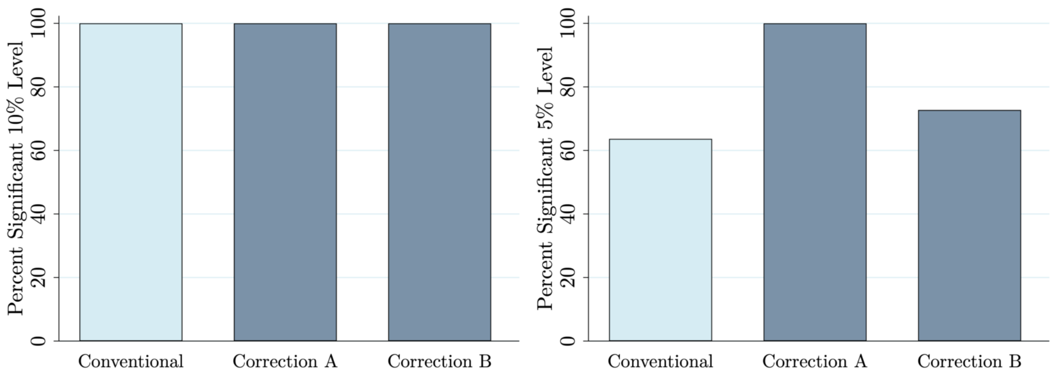
Figure A10. Percent of Results That Remain Significant Under True Cutoff (Zero) and Placebo Cutoffs. Light blue indicates the true cutoff (zero), dark blue indicates alternative cutoffs to the left and to the right of the true cutoff. Refers to Table A28.



(a) 10% Significance

(b) 5% Significance

Figure A11. Percent of Results That Remain Significant Under Default (No Adjustment) and Adjustments (Multiple Hypothesis Testing and Spatial Correlation). Light blue indicates the default (no adjustment), dark blue indicates standard error adjustments. Column 2 shows adjustments to multiple hypothesis testing (false discovery rate), Columns 3 and 4 to spatial correlation (Conley standard errors, 100km, and 500km). Refers to Table A29.



(a) 10% Significance

(b) 5% Significance

Figure A12. Percent of Results That Remain Significant Under Conventional Specifications and Corrections. Light blue indicates the conventional RD estimator with the conventional variance estimator. Correction A is the bias-corrected RD estimator with the conventional variance estimator. Correction B is the bias-corrected RD estimator with the robust variance estimator (Calonico et al., 2014). Refers to Table A25 and A26.

Discussions

Discussion A1. One potential threat is that the IHDS may have been significantly more likely to be conducted in treatment districts. To determine if this is the case, I specify an indicator variable that is 1 if the survey was conducted in a given district and 0 otherwise. I run the main regression specification (Equations 2 and 3) without state-level fixed effects. The resulting coefficient is statistically insignificant (coefficient: 0.07, standard error: 0.20); thus, the survey is not significantly more likely to have been conducted in treatment districts than in control districts. There is no need to conduct this exercise for the DHS or the Economic Census data, since data for all districts is collected.

Discussion A2. I demonstrate that other policies do not pose a threat to identification. The concern is that I may mistake discontinuities around the cutoff for the effect of the 2005 RBI policy when they stem from other policies. To my knowledge, no other policy uses the same cutoff rule described in this paper. For other nationwide policies to coincidentally threaten identification, they would need to be significantly more likely to be implemented in this study's treatment districts. Otherwise, their impact would be smooth around the cutoff. While many policies define certain priority districts, these are unlikely to be identical or highly correlated to treatment districts in this setting. The reason is that priority districts are often defined according to the target of the policy, for instance, certain health indicators. Here, I describe other nationwide policies. The Ministry of Health and Family

Welfare is a government agency that implements health-related policies. In 2005, the ministry initiated the National Rural Health Mission (NRHM). In 2013, the NRHM was joined by the National Urban Health Mission (NUHM), and both approaches were combined under one umbrella, the National Health Mission (NHM). Through these programs, both of which comprise multiple initiatives, the Ministry of Health and Family Welfare focuses on improving health outcomes, especially by targeting the supply of healthcare services. For example, the NRHM includes a safe motherhood intervention scheme that provides cash assistance to promote institutional delivery. Many of these initiatives focus on certain priority states; as there is no variation on the district level, they do not threaten identification. However, in 2013, the ministry published a list of priority districts, which multiple initiatives used as guidance to allocate resources. Priority districts were those that were, within a state, in the bottom quarter of the distribution of a composite health index. For districts with left-wing extremism or a high share of tribal population, those falling in the bottom half of the distribution within a state were included. Because it was implemented in 2013, this definition of priority districts is unlikely to drive the IHDS II findings but could potentially impact health outcomes in the DHS. I do not find any evidence that this is the case. The regression coefficient is insignificant. Additionally, the correlation coefficient within the bandwidth is low at 0.04.

Another ministry that introduced health-related policies is the Ministry of Women and Child Development. One policy in particular is worth considering in this context: the Integrated Child Development Services (ICDS) program. The ICDS was introduced in 1975 and has, among other goals, the objective of reducing mortality, morbidity, and malnutrition. Services under this program include, for instance, immunization and supplementary nutrition. In 2012/2013, a restructured and strengthened ICDS program was rolled out in priority districts. In 2013/2014, a second rollout wave followed. Priority districts were defined based on the nutritional status of children and anemia level among pregnant women. Only the list of the districts in the first wave is available. The regression coefficient is insignificant. The correlation coefficient within the bandwidth is low at 0.05.

Another often discussed nationwide policy is the National Rural Employment Guarantee Act (NREGA) from 2005. It is an employment scheme that guarantees a minimum amount of wage employment for unskilled labor. NREGA was rolled out in three waves. The first was conducted in 2006/2007, followed by one in 2007/2008, and a final wave in 2008/2009. The phase in which each district was covered was based on an index consisting of parameters such as poverty, education, and health. Both the first and the second wave of NREGA have statistically insignificant coefficients. Correlation coefficients are low, with 0.22 and -0.07, respectively.

Finally, I investigate the Rashtriya Swasthya Bima Yojana (RSBY), a national health insurance scheme introduced in 2008 by the Ministry of Labour and Employment. The scheme aims to provide health insurance coverage to Indian citizens who belong to the below-poverty-line population. It provides cashless insurance for hospitalization in public and private hospitals. The scheme was rolled out in priority districts. Similar to the other policies, the RSBY was not more likely to be implemented in the treatment districts of this study; the regression coefficient is insignificant. Additionally, the correlation coefficient is low at -0.08.