Bank Presence and Health

By Kim Fe Cramer*

October 2021

I use a policy of the Reserve Bank of India from 2005 that incentivizes banks to set up new branches in underbanked districts, defined as having a population-to-branch ratio larger than the national average. In a regression discontinuity design, I compare households in districts just above and just below the national average. Six years after the policy introduction, households in treatment districts are a third less likely to be affected by an illness in a month. They miss fewer days of work or school due to an illness and have lower medical expenses. Ten years after the policy was introduced, I observe persistently lower morbidity rates, higher vaccination rates, and lower risks associated with pregnancies. I provide evidence that two previously understudied aspects of banking played a role: households gain access to health insurance and health care providers gain access to credit. In equilibrium, I observe an increase in health care demand and supply.

Keywords: Financial Development, Banks, Health, Insurance, Credit Access
JEL Codes: G21, O12, O16, I15

*Columbia Business School, Columbia University, 3022 Broadway, New York, NY 10027, USA (email: kfc2118@columbia.edu, www.kimfecramer.com). Supported by the Chazen Institute for Global Business and the Bernstein Center for Leadership and Ethics at Columbia Business School. I would like thank Miriam Bruhn, Xavier Giroud, Jonas Hjort, Michaela Pagel, Suresh Sundaresan, Eric Verhoogen, Jack Willis, and Daniel Wolfenzon, as well as numerous seminar participants for their valuable comments. In particular, I would like to thank the Finance and Economics Faculty and Administrative Staff of Columbia Business School as well as the Development Group of the Economics Department for their support. I also would like to thank Apurav Yash Bhatiya from University of Warwick and Yash Pratap Bhatiya from the Oriental Bank of Commerce. Additionally, I am grateful to International House New York for an inspiring environment. Finally, I would like to thank my family and friends for their support.

I. Introduction

What are determinants of poor health status in developing countries? First, households have limited resources to invest in health and limited ability to move available resources across time and states (Dupas and Miguel, 2017). Second, even if households can afford health care, supply is often scarce and low quality (Das and Hammer, 2014). Against this background, what approaches could be effective for improving health outcomes? One potential solution, which I examine here, is developing a country's financial sector. In particular, countries have taken steps to improve bank presence in underserved regions. Households are potentially able to invest more in health if they benefit from increased income as banks stimulate business activity. Households may also be able to invest at the right time if they gain access to savings accounts, credit products, or health insurance. Finally, households could benefit from improved health care supply if banks provide credit constrained health care providers with access to loans. Despite these strong motivations, we lack causal empirical evidence on the impact of bank presence on health.

I use a Reserve Bank of India (RBI) policy, initiated in 2005, that incentivizes banks to set up new branches in underbanked districts, defined as having a population-to-branch ratio that exceeds the national average. In a regression discontinuity design, I initially compare bank presence in districts that have a ratio just above and just below the national average. I then compare the health status of households in these two sets of districts. To measure health, I use two complementary and nationally representative surveys. The first, the Indian Human Development Survey (IHDS), conducted six years after the policy, allows me to measure not only health but also related economic variables such as medical expenses. The second, the Demographic and Health Survey (DHS), was conducted ten years after the policy and has a large sample size that facilitates capturing effects on low-probability events such as miscarriages. With this empirical strategy, I am able to identify the causal impact of bank presence on health.

The foundation of this study lies in its first finding: the RBI policy introduces exogenous variation in bank presence. Before the policy, the number of licenses issued by the RBI and the actual number of branches are smooth around the policy's cutoff. One year after the policy's implementation, significantly more branch licenses are issued for treatment districts just above the cutoff than for control districts just below the cutoff. After one more year, there are significantly more branches in treatment districts. The discontinuities in the number of licenses and number of branches increase over the subsequent years. Five years after the policy was introduced, treatment districts have 21 percent more licenses and 19 percent more bank branches than control districts, an increase to 8.31 branches per 100,000 people, compared to a control mean of 6.99 branches. Pre-policy smoothness, post-policy discontinuities, and dynamics exactly corresponding to the policy timing provide confidence in the hypothesis that the policy induced exogenous variation in bank presence.

The study's second main finding is that the health of households significantly improves with expanded bank presence. Six years after the policy was introduced, households in treatment districts are 36 percent less likely to experience an illness such as fever, diarrhea, or cough in a given month. This compares to a control mean of 52 percent. The reduction in morbidity rates positively affects health-related economic outcomes of households. They miss half a day of work or school less per month due to illness and incur significantly lower medical expenses. Using the second survey, conducted ten years after the policy introduction, I replicate my finding on morbidity rates and demonstrate that health also improves along other dimensions. Households in treatment districts have higher vaccination rates. They also have lower risks associated with pregnancies: as the probability of institutional delivery increases, the probability of miscarriages and stillbirths decreases.

I provide extensive evidence to reject potential threats to causal identification. First, I show that local governments do not manipulate their treatment status. By construction, manipulation of the population-to-branch ratio is unlikely. In the numerator, the total population is historical data from the 2001 Population Census. In the denominator, the total number of branches is the sum of individual decisions of all banks. Additionally, banks directly report their number of branches to the RBI. In alignment, there is no evidence that more districts are located just above than just below the cutoff. Additionally, there is no evidence that districts just above and just below the cutoff are significantly different 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. Additionally, there is no threat to identification due to migration, which is negligible. Finally, there are no policies that use an identical cutoff or that are significantly more likely to be implemented in treatment districts. Results are robust to different bandwidths and polynomials, and there is little evidence of discontinuities at placebo cutoffs. These tests provide confidence in the causal interpretation of my findings.

How exactly does bank presence improve health? To shed light on this question, I lay out a framework of potential mechanisms and discuss which ones have bearing in the data. To motivate my mechanisms, I consider three determinants of poor health status in developing countries. First, households have limited resources to invest in health. Research has shown that bank presence can increase household income (Bruhn and Love, 2014). Indeed, I observe that six years after the policy's implementation, households in treatment districts have 8 percent higher total consumption than those in control districts. I also find that households spend significantly more on food, corresponding to a quarter of a meal more per day. More or better food alone may have had a positive impact on households' health. It is also possible that households increased their health care demand. While there are challenges with observing historical health care demand, increased vaccination rates and higher probability of institutional delivery point towards an increase coinciding with the policy. Summarizing, it is likely that this income channel played a role in improving health; interestingly however, it is complemented by other mechanisms related to previously understudied aspects of banking.

A second important determinant of poor health status is that households have limited ability to shift available resources across time and states. Households might want to prepare for an illness (with a savings account or health insurance) or react to a disease (with credit). I test whether households gain access to these three financial instruments with expanding bank presence.[1] Indeed, I find that six years after the policy, households are significantly more likely to have a savings account and health insurance. The coef-

[1] In the context of developing countries, banks often offer health insurance products themselves or partner with insurance companies to provide access to and verify the identity of potential customers.

ficients on bank loans are positive but insignificant. The literature shows that savings accounts alone are unlikely to drive major welfare changes (Dupas et al., 2018). In contrast, while evidence is scarce, studies suggest that health insurance can positively impact health status (Erlangga et al., 2019). Thus, an increase in health insurance coverage, a largely understudied aspect of bank presence, is likely an important factor in improving health status.

Finally, a third important determinant of poor health status is households' limited access to good health services. Initially, I argue that health care providers generally rely on bank loans and are likely to be credit constrained. Then, I utilize the Economic Census to test whether health care providers gain access to credit due to the policy and whether health care supply increases. Eight years after the policy, hospitals in treatment districts are significantly more likely to report institutional loans as their main source of finance. Additionally, treatment districts have 140 percent more hospitals than control districts eight years after the policy. As expected, there are no discontinuities before the policy, and the reaction is stronger for private hospitals. Household survey evidence supports the hypothesis that supply in treatment districts improved. However, the observed increase in equilibrium health care supply could also result from increased aggregate health care demand. While I do not have exogenous variation in credit to health care providers alone, my findings are in alignment with the mechanism of banks providing credit access to health care providers, allowing them to improve supply. Credit access to providers of essential services to households has been previously understudied in the literature.

Insofar as the relationship between financial development and health has been studied, researchers have used randomized controlled trials (RCTs) to provide households with access to financial instruments. These studies investigate a wide range of outcomes, including health. A recent wave of papers identifies the impact of savings accounts (Dupas and Robinson, 2013; Dupas et al., 2018; Prina, 2015), bank credit (Karlan and Zinman, 2010), and microcredit (Angelucci et al., 2015; Augsburg et al., 2015; Attanasio et al., 2015; Banerjee et al., 2015a; Crépon et al., 2015; Tarozzi et al., 2015). They consistently find null effects on health. Few RCTs evaluate health insurance and most stop short at measuring insurance take-up or health care utilization. Those that examine the impact on health status partly find positive effects (Erlangga et al., 2019). This paper makes three important contributions to this literature. First, using a natural experiment allows me to study access to financial services for three types of agents: households, businesses, and health care providers. Second, for each type, I examine access by many agents. Finally, I observe long-term effects up to ten years after the policy introduction; most RCTs only consider effects over three years. To summarize, my setting allows me to realistically capture effects that emerge due to interactions between financial access by different type of agents, and due to a higher diffusion or over a longer time frame.

More broadly, this paper speaks to the question of how financial development affects household wellbeing. RCTs providing households with financial access, discussed above, can be viewed as one strand in this literature. In a second strand, researchers use natural experiments, employing variation in bank presence or microcredit availability. Studies have shown positive effects on income, consumption, poverty, and financial resilience (Bruhn and Love, 2014; Breza and Kinnan, 2021; Burgess and Pande, 2005; Bharadwaj et al., 2019; Célerier and Matray, 2019). The first contribution of this paper to this literature strand is that it extends the impact evaluation to another important dimension of households' wellbeing: health. Critically, we cannot simply extrapolate from the findings in earlier studies to a positive impact on health. Even as households become richer, factors such as lack of information, non-monetary transaction costs, or behavioral biases might mean that they do not invest significantly more in health (Dupas and Miguel, 2017). Additionally, even if households increase their health care demand, this might not translate into significant improvements in health if the supply of services remains low. In particular, a gradual increase in demand might not stimulate an increase in supply if there are large fixed costs to investments. Reinforcing the claim that previous findings cannot simply be extrapolated to a positive effect on health, studies that provide exogenous income variation in form of cash transfers find no short- or long-term effects on health (Haushofer and Shapiro, 2013, 2018; Egger et al., 2019). The second main contribution of the paper is that it puts two aspects of banking into focus that previous literature has taken little notice of: access to health insurance for households and access to credit for providers of services crucial for households. Earlier studies have focused on credit access to businesses or access to savings accounts or credit products for households.

Finally, this paper relates to the comprehensive literature on the relationship between financial development and economic growth. Moving from cross-country studies (Goldsmith, 1969; King and Levine, 1993) to industry- and firm-level studies (Rajan and Zingales, 1996; Beck et al., 2005), this literature has now largely established that development of the financial sector can positively influence economic growth. However, empirical studies have mainly focused on the mechanism of increased access to credit for businesses (Levine, 2005). This paper contributes evidence of another potential mechanism: that financial development can improve health status of citizens, resulting in increased labor supply and school attendance. Both of these aspects are not only beneficial for the household itself, but may also positively influence economic growth on an aggregate level.

My findings have important implications for policy as well as for future research. Policymakers can conclude that incentivizing bank presence can be beneficial for households' health. Since one mechanism appears to be the interaction between financial service providers and health care providers, policymakers might want to concentrate on incentivizing the relationship between financial actors and providers of services they want to foster.[2] This paper also speaks to researchers, suggesting promising new areas of interest. One open question is to what extent the different mechanisms contribute to improving health. Answering this requires exogenous variation in respective channels, for instance, credit access to health care providers only. A second line of inquiry is whether other dimensions of wellbeing, such as education, show positive impact when evaluated in a context of a natural experiment. For instance, providers of educational services might benefit from being provided credit access. Gaining an understanding of these questions could significantly advance our knowledge of the impact of banks and the scope of policymakers to improve their citizens' wellbeing.

[2]This is indeed what the RBI did in light of the recent COVID-19 crisis. The reserve bank announced a policy in May 2021 to incentivize banks to quickly deliver credit to health care providers, injecting USD 6.78 billion of liquidity.

The rest of this article is organized as follows. Section II introduces the policy that gives rise to the natural experiment. Section III outlines the data I use in my analysis. In Section IV, I explain the identification strategy and provide empirical evidence that the identification assumption holds. Section V shows that bank presence increases in treatment districts due to the policy. In Section VI, I demonstrate that health status of households improves. This is followed by a detailed discussion on mechanisms in Section VII. Section VIII discusses robustness and placebo tests. I conclude with Section IX.

II. 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. RBI's definition of an underbanked district is crucial for identification in this study: they are districts with a population-to-branch ratio that 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 III. 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 for 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 reflects in the dynamic patterns of my analysis on bank presence.

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

Figure 1 depicts all 593 districts as of the 2001 Census. Marked in dark blue are the 375 districts that are defined as underbanked according to the reconstructed district-level ratio in 2006.

To my knowledge, this is the first paper that combines the 2005 RBI policy with household-level data. Other authors that use this policy, such as Young (2017), focus on aggregate outcomes and different questions such as whether bank presence fosters economic activity. A similar branch licensing policy was in place between 1977 and 1990. Burgess and Pande (2005) use that policy in their seminal paper on the impact of bank presence on poverty, employing a different empirical strategy than mine. The authors focus on state-level measures of poverty, and exactly how it is reduced remains in a black box. Here, I provide empirical evidence of one potential mechanism for poverty reduction: improved health. From 1990 through 2005, no comparable branch licensing policy was in place.

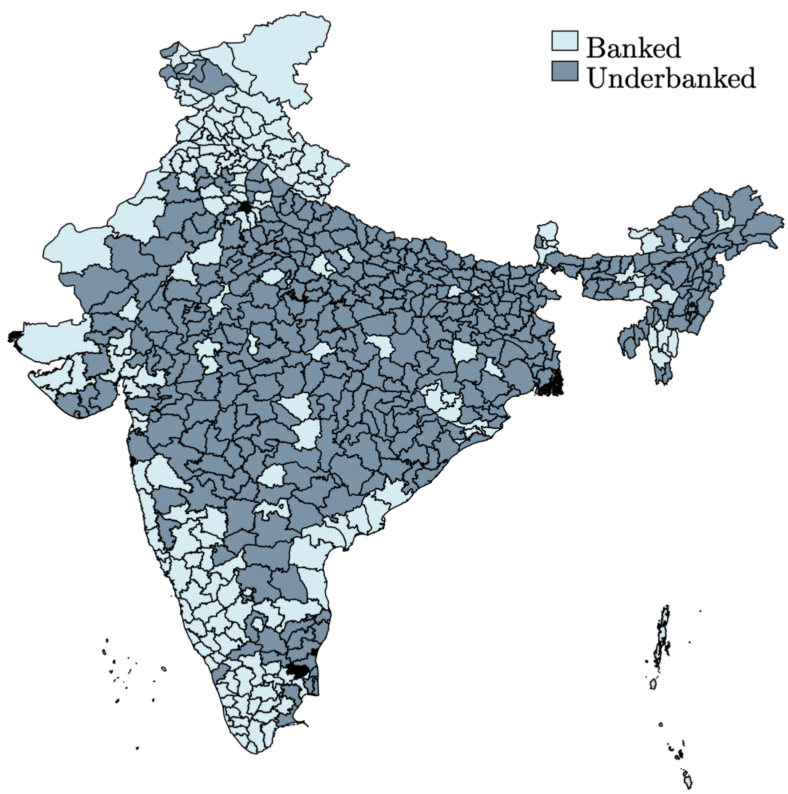


Figure 1. Banked and Underbanked Districts. District borders refer to the 2001 Census.

III. Data

Initially, I reconstruct the policy's population-to-branch ratio. To measure population of each district, I rely on the 2001 Population Census. To obtain the number of branches in a given district, I use the Master Office File published by the RBI. Because the final list of underbanked districts was issued in 2006, I refer to the first quarter in this year for my measure. The Master Office File also allows me to study the dynamic reaction of banks to the policy. One specific type of bank, regional rural banks, is excluded from the policy; correspondingly, I also exclude this bank type from my variables 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 as well as 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. I do not use data from after 2016, the year in which the final household-level survey was conducted. General summary statistics from the Master Office File are provided in 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). The pre-policy round allows me to test for smoothness of household characteristics around the cutoff. The post-policy round provides the main outcome variables. Importantly, the survey not only contains information on health, but also provides a holistic picture of the households' economic situation. With this data, I can test, for instance, whether households hold various financial instruments or how many days of work or school they missed due to illness. The first survey round was conducted in 64 percent of districts, and the second in 65 percent. Figure 1(a) 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 depict a threat to identification, as I rely on a comparison of households in treatment and control districts in the second survey round. General summary statistics of the IHDS are provided in Table A2. I also provide evidence on 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 along dimensions of consumption, financial access, and health.

I complement the IHDS with a second nationally representative household-level survey, the Demographics and Health Program (DHS),[3] conducted in 2015/2016, ten years after the policy (see Figure 2). In contrast to the IHDS, the DHS primarily focuses on health. The advantage of this survey is that it has a very large sample size, which allows me to capture effects on low-probability events such as miscarriages. The DHS was conducted in all districts, as indicated in Figure 1(b), and interviewed 601,509 households. The previous round of this survey, conducted in 2005/2006, 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.

To examine the mechanisms driving the relationship between bank presence and health, it becomes crucial to understand the response of health care supply. For this purpose, I primarily rely on the Economic Census, from which I obtain measures of the number of hospitals, other medical service providers, and general businesses, as well as information about the major source of finance for these establishments. I focus on two census rounds; the first was conducted in 2005 and the second in 2013 (see Figure 2). The first Economic Census round allows me to test for smoothness around the cutoff in the respective variables pre-policy. The second round provides outcomes variables. General summary statistics are provided in Table A4. To generally gain a better understanding of the health care sector, I investigate summary statistics from the Prowess database, which provides financial statements for companies of all sizes, including those conducting health services. While providing more detailed financial information than the Economic Census, the Prowess

[3]The survey is also known under the name National Family Health Survey (NFHS).

database only contains a selective sample of health care providers for a limited number of districts. For my regression analysis, I thus concentrate on the Economic Census.

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) (Asher et al., 2021; Henderson et al., 2011). This platform combines multiple data sources on the village or town level. Economic activity is proxied by night-light data, economic employment, and road connection. Population characteristics include total population, rural and urban 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. 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 exercise is the 2011 Census.

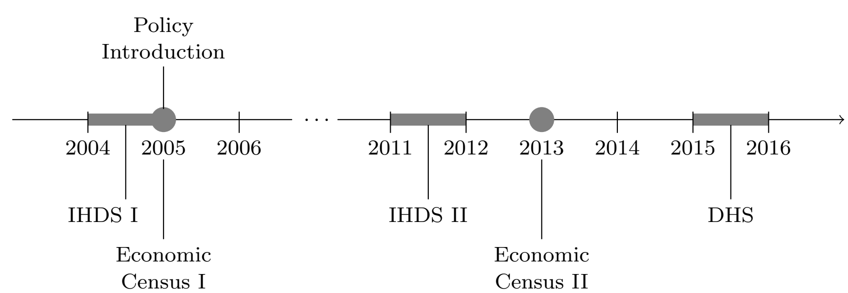
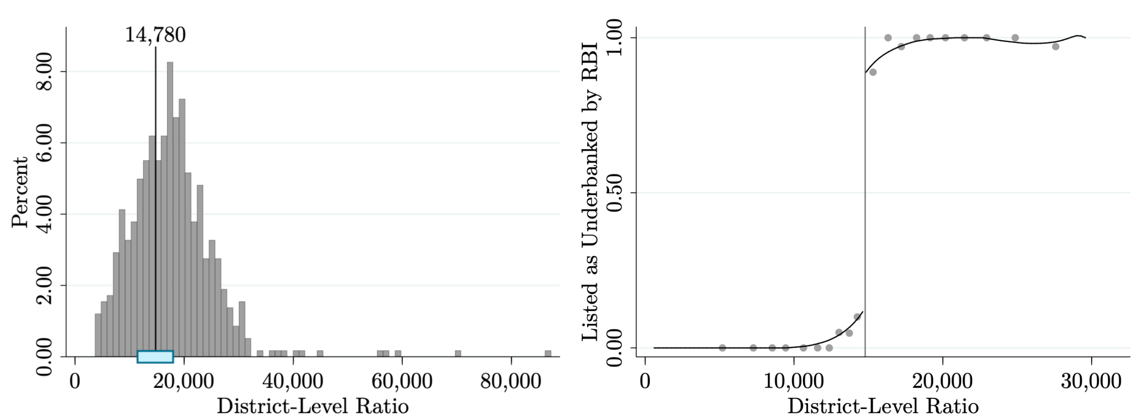


Figure 2. Timeline. The following graphic depicts a timeline of this study, with the three main data sets used (IHDS, DHS, Economic Census).

IV. 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 with a ratio 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, districts included are mostly within a range of $\pm 3,000$ relative to the cutoff, indicated by the blue bar on the x-axis in Figure 3(a). 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. On a first look, 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.



(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).

While I do not perfectly predict which districts are listed as underbanked by the RBI, there are few districts, 10 out of 593, that have a different status than predicted. Figure 3(b) shows that when a district's ratio crosses the national average, there is a large jump in 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.

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

$$\text{(3)} \quad \begin{aligned} \text{y}_{h,d,s} &= \beta_0 + \beta_1 \text{Underbanked}_{d,s} + \beta_2 \text{DistRatio}_{d,s} \\ &+ \beta_3 \text{DistRatio}_{d,s} \text{Above}_{d,s} + \gamma X_{d,s} + \eta_s + \epsilon_{h,d,s} \end{aligned}$$

Here h denotes household, d denotes district, and s denotes state. $\text{Underbanked}_{d,s}$ is an indicator that is one if the district is listed as underbanked. $\text{DistRatio}_{d,s}$ is the district-level ratio. $\text{Above}_{d,s}$ is an indicator that is 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 continuity of all characteristics other than being underbanked at the cutoff. This assumption is violated if agents precisely manipulate the ratio of their district. To understand how systematic differences could be introduced by manipulation, consider the following. Assume that local governments hear about the policy and want to benefit from more banks in their area. Also assume that they have the ability to manipulate the population-to-branch ratio, such that they move from being just below the cutoff to just above the cutoff. If these districts have a particularly rich or healthy population, I will confuse their characteristics with a treatment effect of the policy.

However, manipulation of the population-to-branch ratio is unlikely due to its construction. First, the numerator contains population data from the 2001 Census. In order to manipulate this historical data, local governments would have to have anticipated the detailed policy rule years prior to its implementation. Second, the denominator is the sum of individual decisions of all banks. 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. Additionally, banks directly report their number of branches to the RBI, leaving no room for an intermediary party to manipulate.[4] Additionally, manipulation has two implications that are empirically testable.

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. On a first glance, there is no evidence of this in 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 already be different from districts just below the cutoff before the policy. Assume, for example, that local governments that are able to manipulate their ratio have a richer or healthier population. In this case, I would observe discontinuities in pre-policy consumption and health measures just around the cutoff. To test whether pre-policy characteristics are smooth, I initially utilize the IHDS I, conducted in 2004/2005. Table 1 demonstrates that households have

[4] Note that any discretion on the level of the RBI is not related to manipulation, but gives rise to the fuzzy RDD. For instance, the RBI might have decided to include a district in the list despite it having a ratio below the cutoff, resulting in extra banks to the left of the cutoff. This is reflected in the first stage and does not pose a threat to causal identification.

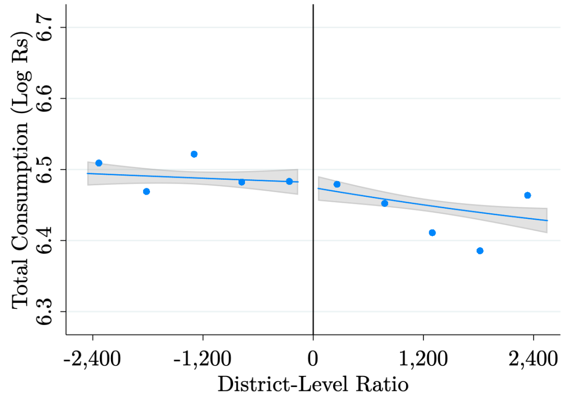
smooth consumption, financial access, and health status before the policy. Columns 1 and 2 show the mean for all treated and control households. Columns 3 and 4 depict means only for observations within the optimal bandwidth. Column 5 reports the fuzzy RDD coefficients, referring to β_1 as defined above. As expected, all coefficients are statistically insignificant. Households in treatment districts do not have higher consumption, higher financial access, or better health status ex ante. Correspondingly, I observe graphical smoothness in Figure 4.

I complement the evidence of smoothness of household variables with smoothness of district characteristics. I demonstrate that districts have smooth bank presence in Section V and that they have smooth hospital presence in Section VII. To obtain information on bank presence, I rely on dynamic data of the RBI. To measure hospital presence, I utilize the Economic Census conducted in 2005. Additionally, I use the SHRUG to provide village- or town-level evidence of smoothness of general economic activity and population characteristics. Smoothness of respective characteristics is indicated in Table A5. Taken together, these tests do not provide any evidence of manipulation, thus they strengthen confidence in the causal interpretation of the RDD estimates.

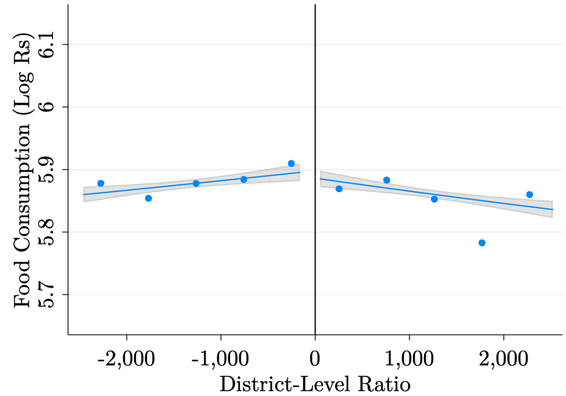
Table 1: Smooth Pre-Policy Covariates

	All observations		Within bandwidth		RDD
	Treated	Not treated	Treated	Not treated	Coefficient
	(1)	(2)	(3)	(4)	(5)
Consumption					
Total consumption (log Rs)	6.38	6.57	6.42	6.51	-0.01
	(0.42)	(0.42)	(0.43)	(0.42)	(0.05)
Food consumption (log Rs)	5.81	5.95	5.84	5.90	-0.03
	(0.32)	(0.32)	(0.33)	(0.32)	(0.03)
Financial Access					
Any loan (yes/no)	0.50	0.42	0.50	0.45	0.00
	(0.50)	(0.49)	(0.50)	(0.50)	(0.10)
Largest loan from bank (yes/no)	0.11	0.12	0.12	0.12	0.00
	(0.31)	(0.32)	(0.33)	(0.32)	(0.03)
Largest loan amt (log Rs)	3.87	2.38	3.65	3.03	0.12
	(4.46)	(4.08)	(4.47)	(4.35)	(0.86)
Health insurance (yes/no)	0.02	0.04	0.02	0.02	0.01
	(0.14)	(0.18)	(0.15)	(0.15)	(0.01)
Health					
Days ill (yes/no)	0.53	0.40	0.48	0.41	-0.06
	(0.50)	(0.49)	(0.50)	(0.49)	(0.06)
Days ill (log no.)	0.86	0.61	0.75	0.64	-0.11
	(0.97)	(0.89)	(0.94)	(0.90)	(0.13)
Days missed (yes/no)	0.41	0.30	0.33	0.34	-0.11
	(0.49)	(0.46)	(0.47)	(0.48)	(0.08)
Days missed (log no.)	0.58	0.42	0.45	0.48	-0.19
	(0.84)	(0.74)	(0.77)	(0.78)	(0.14)
Treatment spending (yes/no)	0.51	0.39	0.46	0.40	-0.08
	(0.50)	(0.49)	(0.50)	(0.49)	(0.06)
Treatment spending (log Rs)	1.68	1.25	1.57	1.32	-0.14
	(2.26)	(2.11)	(2.22)	(2.15)	(0.27)

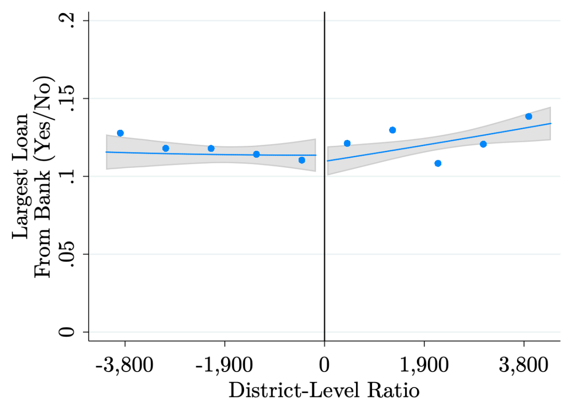
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. Data IHDS I (2005/2006). Household level. All variables that are measured in currency or days are transformed into log form and trimmed at the 10th and 90th percentile. Days missed measures the number of days that the household was not able to do usual activities and had to miss work or school. Consumption measures are indicated per household member.



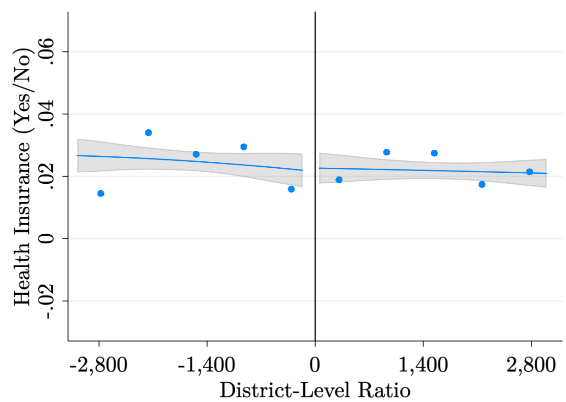
(a) Pre-Policy (2004/2005)



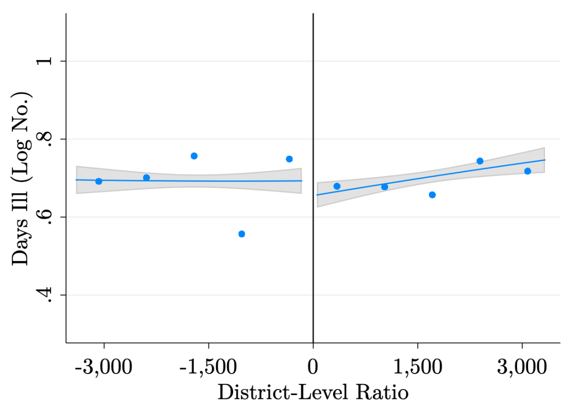
(b) Pre-Policy (2004/2005)



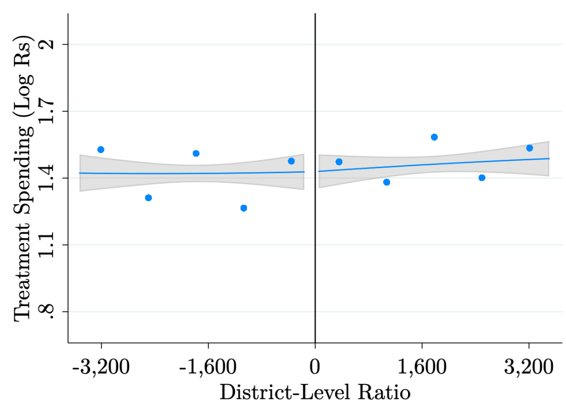
(c) Pre-Policy (2004/2005)



(d) Pre-Policy (2004/2005)



(e) Pre-Policy (2004/2005)



(f) Pre-Policy (2004/2005)

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 90 percent confidence intervals indicated in gray. The cutoff is normalized to zero.

A second potential threat to identification is migration. If households migrate to treatment districts as a result of increased bank presence in these areas and these households are, for instance, richer or healthier, I will 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 prior to the 2011/2012 IHDS II. The coefficient on this migration pattern is insignificant when formally testing for it as described in the regression framework. Table A6 provides evidence that migration does not pose a threat to identification.

Finally, I demonstrate that other policies do not pose a threat to identification. The concern here is that I may mistake discontinuities around the cutoff for the effect of the 2005 RBI policy when they actually stem from other policies. To my knowledge, there is no other policy that uses the same cutoff rule as described in this paper. In order 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 A1, I provide a thorough description of other nationally implemented policies, including those issued by the Ministry of Health and Family Welfare and the Ministry of Women and Childhood Development, 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 certain policy and zero otherwise. Using this indicator variable as an outcome in my main regression specification, I test whether the policy was significantly more likely to be implemented in treatment districts (Table A7). All coefficients are statistically insignificant, as expected. I provide further evidence on the distribution of priority districts in Table A8. Within a bandwidth of $\pm 4,000$, priority districts depict a low share of overall districts, ranging from 19 to 28 percent. Additionally, the difference in percent of priority in treatment and in control districts within the bandwidth ranges from -6 to 23 percentage points. Correlation coefficients between an indicator for priority district and an indicator for being above the cutoff within the bandwidth range from -0.07 to 0.25. This evidence suggests that other policies do not pose a threat to this study's causal identification.

V. Banks Open Branches

In the first step of the analysis, I provide evidence that banks indeed reacted to the policy. I examine two outcomes: number of branch licenses and number of branches. Since I observe years between 1997 to 2016, I test both for smoothness pre-policy and for discontinuities post-policy. In Table 2, I examine the number of branch licenses in 2004, one year before the policy, and in 2010, five years after the policy.[5] As expected, coefficients in the year before the policy are not statistically significant. Treatment districts have neither more branch licenses nor more branches than control districts. Post-policy, as expected, I observe statistically significant discontinuities in both branch licenses and branches. In 2010, treatment districts have on average 21 percent more branch licenses and 19 percent more branches than control districts. The latter corresponds to an increase to 8.31 branches per 100,000 people, compared to the control mean of 6.92 branches. Turning to the graphical analysis, I observe discontinuities in branch licenses (Figure 5(c)) and a discontinuity in branches (Figure 5(d)) five years after the policy.

Table 2: Banks Open Branches

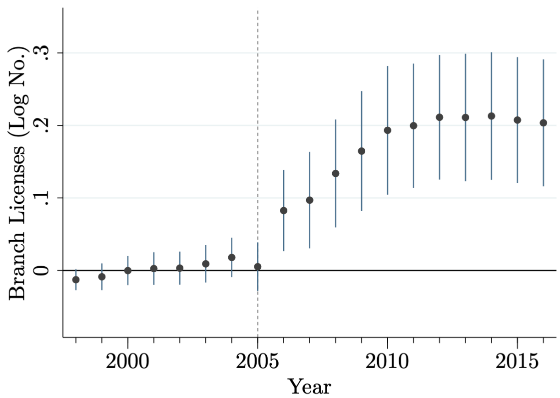
	Pre-Policy (2004)		Post-Policy (2010)	
	Branch Licenses (log no.) (1)	Branches (log no.) (2)	Branch Licenses (log no.) (3)	Branches (log no.) (4)
Treated	0.02	0.01	0.19***	0.17***
	(0.02)	(0.02)	(0.05)	(0.06)
Control Mean	4.00	3.98	4.38	4.38
Change (%)	1.81	1.01	21.32	18.98
First Stage	0.81	0.80	0.80	0.80
Bandwidth	3,490	3,621	2,972	3,329
Efficient Obs.	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 in parentheses. 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.

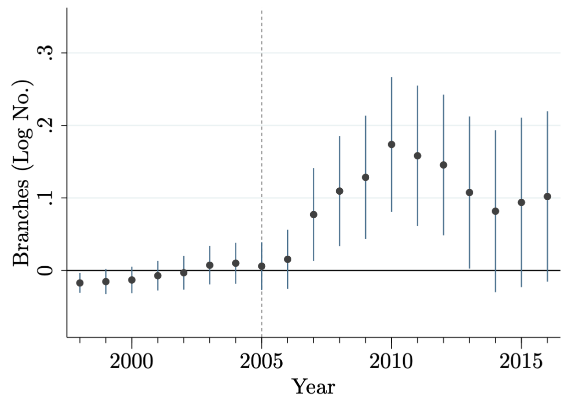
[5] Tables that describe treatment effects contain the following information: The first line provides the main coefficient of interest, β_1. This is followed by the control mean within the optimal bandwidth as well as the mean change. Next, the reader can find the first stage coefficient, α_1. Following that are the optimal bandwidth and the number of efficient 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 any baseline controls are included in the regression.

Importantly, the dynamic response of banks corresponds exactly to the policy timing. Figure 5(a) depicts the dynamics for branch licenses and Figure 5(b) for branches. As expected, there is smoothness around the cutoffs before the policy, and coefficients become statistically significant subsequently to the policy. The reaction in branch licenses issued is immediate: the coefficient on branch licenses becomes statistically significant in 2006, the year that the final list of underbanked districts is published. As expected, the reaction in branches is slightly lagged by one year: the coefficient becomes statistically significant in 2007. There is another interesting pattern that can be explained by the policy. In 2010, as discussed in Section II, 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 number of branches—corresponds exactly to what one expects to see if banks increasingly open branches in districts to the left of the cutoff without a license in some states. 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. Aggregated district-level credit and deposit amounts are discussed in Figure A10, and branch profitability is examined in Discussion A3. One can conclude that banks indeed reacted to the policy, providing exogenous variation in bank presence.

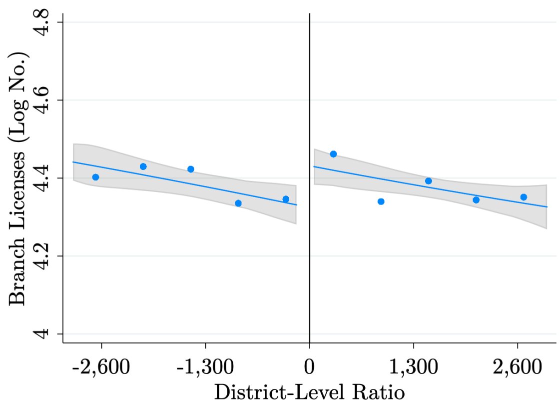
Standard robustness and placebo tests on bank outcomes are discussed in Section VIII, but one placebo test that emerges from the design of the policy is discussed here. One type of bank, regional rural banks, are excluded from the policy. Consequently, one does not expect to observe positive coefficients for this bank type. I test for discontinuities in branch licenses and branches of regional rural banks in 2010 (Table A9), and, as expected, coefficients in the placebo test are insignificant.



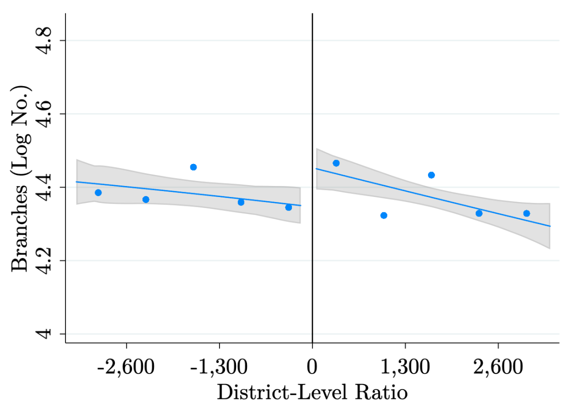
(a) Dynamics



(b) Dynamics



(c) Post-Policy (2010)



(d) Post-Policy (2010)

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

VI. Health

In the second step of the analysis, I investigate the effects of bank presence on three pillars of health— morbidity rates, vaccination rates, and pregnancy risks—and demonstrate that the health status of households improves in treatment districts. For this, I rely on two complementary household surveys. The IHDS II (2011/2012) not only allows me to measure morbidity rates, but also paints a holistic picture of the economic situation of the households. For instance, it allows me to measure how many days of work or education the household missed in the past month due to an illness. The DHS (2015/2016), on the other hand, has a very large sample size, allowing me to capture effects on low-probability events such as miscarriages. Additionally, using a second data set allows me to replicate findings from the first one, providing further confidence in my results.

A. Lower Morbidity Rate

Initially, I investigate the effect of bank presence on morbidity rates, which are measured six years after the policy in the IHDS II. I concentrate on illnesses related to fever, cough, or diarrhea. I set aside chronic diseases such as cancer, which are less likely to be responsive to the policy. In a module on health, households are asked whether any household member was ill in the past month and how many days members were ill in total. Results are depicted in Column 1 and 2 of Table 3. I observe that households in treatment districts have a 36 percent lower probability to have any member affected by fever, diarrhea, or cough. Roughly speaking, every second household in control districts is affected by an illness in a given month, while in treatment districts, only every third household is affected. In alignment, the total number of days that household members are ill is 25 percent lower in treatment districts. Pre-policy smoothness for this variable is depicted in Figure 4(e). The post-policy discontinuity is shown in Figure 6(a). This evidence indicates that the health status of households improves due to the policy.

Illnesses can have important economic consequences for households. First, if someone in the household falls ill, that person is unlikely to be able to go to work or school, resulting in income loss or missed learning experiences for the household. Second, the household has pay for medical expenses. The survey data allows me to capture the effect on these economic consequences of illnesses. Based on the significant improvement in health status, I expect treated households to miss less work or school and spend less on medical expenditures related to fever, cough, or diarrhea. Indeed, this is what I observe in Table 3. Column 3 and 4 refer to the question whether (or how many days) members missed work or school due to an illness. Households in treatment districts are 71 percent less likely to give an affirmative answer than those in control districts. On the internal margin, the number of days missed is 35 percent lower in treatment districts, corresponding to a decrease to 1.05 days compared to a control mean of 1.62 days. In terms of medical expenses, households are asked to report whether (or how much) they spent on treatment of fever, cough, or diarrhea in the past month. The effect on treatment spending is depicted in Column 5 and 6. I find that the probability that households spend on treatment is 34 percent lower in treatment districts. This is a significant improvement compared to the control mean, which shows that more than half of the households spend on treatment. It is also in alignment with the coefficient size in Column 1 on probability of illness. In control households, the amount spent on medical expenses is 121 rupees on average, which means that treatment households save 71 rupees, or around \$1.32 per month. These effects are not calculated conditional on having an illness. Pre-policy smoothness of medical expenses is depicted in Figure 4(f), and post-policy discontinuity in Figure 6(b). I conclude that the improvement in health status is accompanied by a decrease in economic costs borne by households.

Table 3: Lower Morbidity Rate

	Morbidity		Economic consequences			
	Days ill		Days missed		Medical expenses	
	(yes/no)	(log no.)	(yes/no)	(log no.)	(yes/no)	(log Rs)
	(1)	(2)	(3)	(4)	(5)	(6)
Treated	-0.19**	-0.29**	-0.30***	-0.44***	-0.18**	-0.88**
	(0.09)	(0.12)	(0.10)	(0.13)	(0.08)	(0.35)
Control Mean	0.53	0.82	0.41	0.58	0.52	2.12
Change (%)	-35.74	-25.21	-71.46	-35.40	-33.61	-58.56
First Stage	0.65	0.70	0.66	0.68	0.66	0.69
Bandwidth	2,204	2,658	2,331	2,513	2,373	2,948
Efficient Obs.	11,749	12,968	12,730	12,421	12,862	14,576
Observations	35,103	32,280	36,805	33,346	36,805	32,983
Baseline Control	No	No	No	No	No	No

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. Data IHDS II (2011/2012). Household level. All variables measured in currency Rs are in log form and trimmed at the 10th and 90th percentile. All illnesses refer to fever, diarrhea, or cough. Days missed measures the number of days that the household was not able to do usual activities and had to miss work or school. All questions refer to the past 30 days.

All variables considered in this analysis are balanced on baseline as demonstrated in Table 1 and Figure 4(e) and 4(f). The full regressions with pre-policy variables are depicted in Table A23. Additionally, Table A24 repeats the analysis including baseline measures of the outcome variable as control. The coefficients remain statistically significant. Since long-term illnesses such as cancer, diabetes, or heart diseases are less likely to be affected in probability, they depict a natural placebo outcome. I repeat the analysis for these kind of diseases in Table A25. As expected, coefficients are insignificant.

B. Higher Vaccination Rate

Increasing child vaccination rates is another driver of health improvements and is crucial for reducing child mortality. Usually, children are vaccinated against tuberculosis, diphtheria, whooping cough, tetanus, polio, and measles. Due to the policy, the fraction of households with children that have received any vaccination is 8 percent higher in treatment districts. This corresponds to an increase to 93 percent of households, compared to a control mean of 86 percent (Table 4, Column 1). The discontinuity is graphically depicted in 6(c). The DHS measures vaccination rates for all children below five approximately ten years after the policy, in 2015/2016. Measuring the effect on vaccination provides further confidence in my results, since the results are less likely to be affected by self-reporting biases. Over half of the affirmative answers on vaccination status were obtained from an official vaccination card. In addition to examining the effect on vaccination rates, I utilize the data to replicate my findings on morbidity rates. Again, I find a negative effect on the probability of illnesses such as fever, diarrhea, or cough. The DHS collects this data only for children below five, not for other household members. Column 2 of Table 4 indicates that the fraction of households with a child that fell ill in the past two weeks is 23 percent lower in treatment districts, corresponding to a decrease to 21 percent in treatment districts compared to a control mean of 27 percent. The discontinuity is observable in 6(d). Another proxy for health status is the number of visits to health care providers. This is an imperfect proxy, since it could also reflect households shying away from visits. I do find that households in treatment districts are 26 percent less likely to have gone to a health care provider for any reason in the past three months. In particular, households are 23 percent less likely to have gone to a provider for treatment of a child. This allows for the conclusion that children not only have a higher probability to be vaccinated, they also benefit generally from better health.

Table 4: Higher Vaccination Rate

	Vaccination	Morbidity	Health care visits	
	Vaccinated child (yes/no) (1)	Sick child (yes/no) (2)	Any reason (yes/no) (3)	Children's treatment (yes/no) (4)
Treated	0.07*	-0.06*	-0.08**	-0.02*
	(0.04)	(0.03)	(0.03)	(0.01)
Control Mean	0.86	0.27	0.29	0.11
Change (%)	8.34	-23.12	-26.84	-22.99
First Stage	0.72	0.70	0.72	0.73
Bandwidth	2,898	3,539	3,287	3,383
Efficient Obs.	26,117	66,658	166,756	187,208
Observations	86,079	171,471	431,148	471,985
Baseline Control	No	No	No	No

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. Data DHS (2015/2016). Household level. Data on health status is coded as missing for households without children below five. Data on health care visits is obtained from the women's questionnaire; households without an eligible woman are coded as missing.

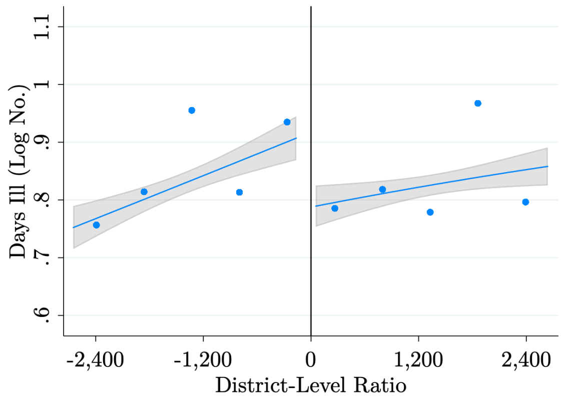
C. Safer Pregnancies

Another important aspect in which I observe improvement is maternal health. There are still significant risks associated with pregnancies in developing countries. In this section, I demonstrate that the policy played an important role in making progress on this dimension. Table 5 depicts the effect of the policy on respective outcomes, measured in the DHS (2015/2016). Initially, I observe that households in treatment districts are 34 percent more likely to have a woman who reported to have delivered in a health care facility (Column 1). The discontinuity is depicted in Figure 6(e). The low control mean arises because the question only refers to pregnancies in the past three months and is coded zero if there was no pregnancy. In alignment with this observation and potentially as a result of it, I find that the fraction of households with a woman who ever experienced a miscarriage or stillbirth is significantly lower in treatment districts (Column 2 and 3). The fraction of households with a woman who experienced a miscarriage is 26 percent lower, corresponding to a decrease to 2.80 percent compared to a control mean of 3.78 percent. The discontinuity is graphically observable in Figure 6(f). The fraction of stillbirths decreases from a control mean of 0.45 percent to 0.24 percent, a reduction of 46 percent. While miscarriages and stillbirths are low-probability events, they can have large consequences for women's physical and mental health. Consequently, any progress on this dimension is of high importance. Finally, I provide indirect evidence that effects on general morbidity rates are also replicable for women in the 2015/2016 survey. While data on health status related to, for example, fever, cough, or diarrhea is only collected for children, I do observe whether women went to a health care facility in the past three month for their own treatment. The fraction of households with women who went for a visit is 30 percent lower in treatment than control districts. Summarizing this evidence, women's health appears to improve as a result of the policy.

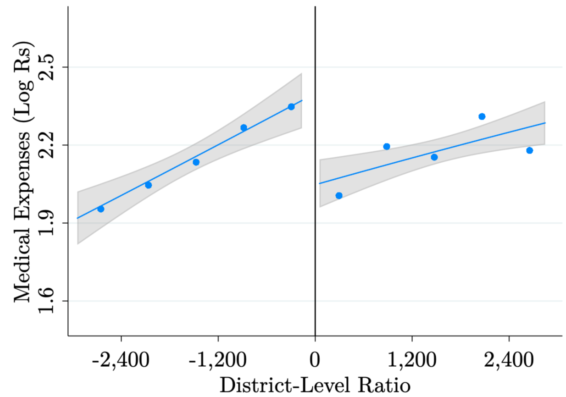
Table 5: Safer Pregnancies

	Pregnancies			Visits
	Health care facility delivery (yes/no) (2)	Experienced miscarriage (yes/no) (2)	Experienced stillbirth (yes/no) (3)	Women's treatment (yes/no) (4)
Treated	0.005***	-0.010*	-0.002*	-0.051*
	(0.002)	(0.006)	(0.001)	(0.027)
Control Mean	0.016	0.038	0.004	0.170
Change (%)	33.52	-26.30	-45.92	-29.84
First Stage	0.72	0.73	0.73	0.72
Bandwidth	3,023	3,430	3,386	3,277
Efficient Obs.	172,892	188,571	187,208	182,318
Observations	471,985	471,985	471,985	471,985
Baseline Control	No	No	No	No

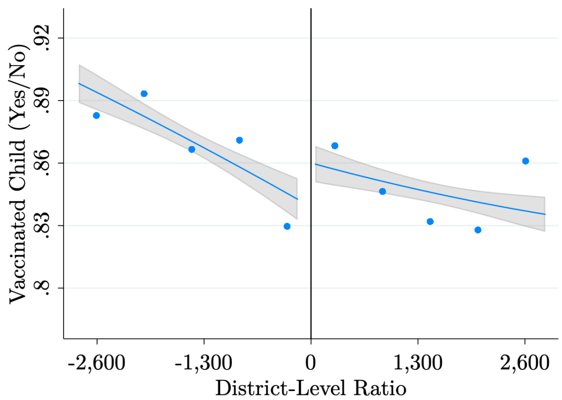
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. Data DHS (2015/2016). Household level. Data on health status and health care visits is coded as missing for households without an eligible woman.



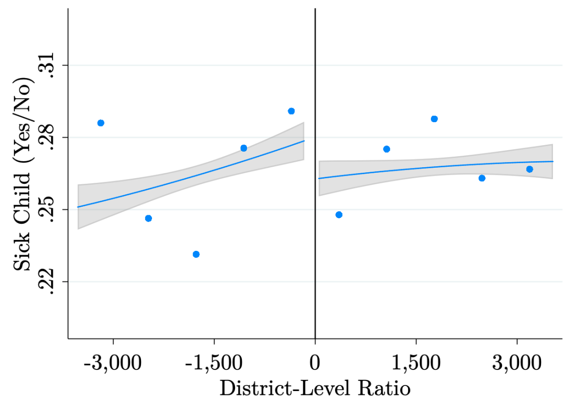
(a) Post-Policy (2011/2012)



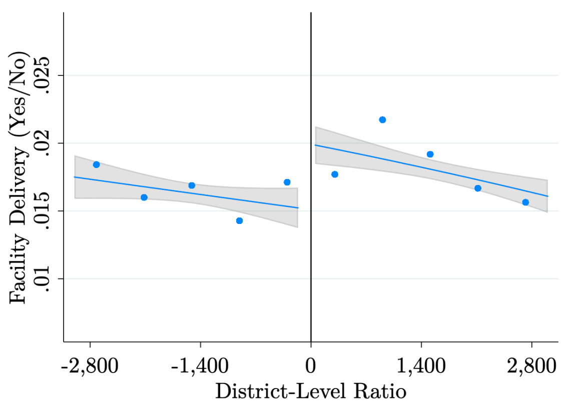
(b) Post-Policy (2011/2012)



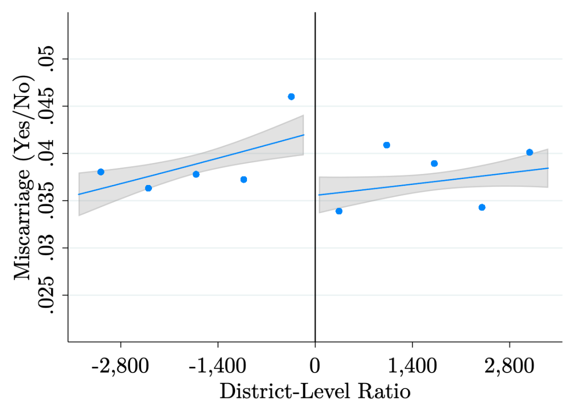
(c) Post-Policy (2015/2016)



(d) Post-Policy (2015/2016)



(e) Post-Policy (2015/2016)



(f) 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 90 percent confidence intervals indicated in gray.

D. Discussion of Effect Sizes

To provide a further piece of evidence of the validity of my findings, I compare my effect sizes to those of other successful health interventions. The study that measures an outcome most closely to my context is Gertler (2004), which evaluates a large conditional cash-transfer program, Progresa, initiated in Mexico in 1977. Eligible families received a cash transfer of about 25 percent of household income every two months. The eligibility conditions were designed to improve health status of families. The payout was received if, for instance, children got immunized and mothers visited nutrition monitoring clinics. In 1998, the government enrolled entire villages, but randomly chose which ones to enter initially and which to enter two years later. Using this experimental design, the author compares households in treatment villages to those in control villages, two years after Progresa was rolled out. Very similar to the morbidity measure for children in DHS, the author uses as an outcome variable a question regarding whether the child was ill in the past four weeks. The study finds that after two years of program exposure, children age zero to three at baseline (or two to five at time of the survey) experience a decrease in probability of being ill by 39 percent. In contrast, I find that children who lived in a treatment district for ten years have a 27 percent lower probability of illness. Here, I directly run child-level regressions to make the results as comparable as possible, but some differences remain. For example, the estimate of the conditional cash transfer refers to the past four weeks, while the DHS refers to the past two weeks. However, one can make the cautious statement that ten years of bank exposure result in an effect on the probability of illness among children that is approximately 70 percent as large as two years of exposure to a program that includes large cash transfers of around 25 percent of household income and that directly incentivizes families to engage in behavior designed to improve health.

Other successful health interventions provide similarly large effect sizes. Examining the effect of improved water quality, Kremer et al. (2011) find that children's probability of diarrhea decreases by 25 percent. Even stronger effects are observed in successful supply-side interventions. Evaluating the impact of trained informal health care providers, Björkman-Nykvist et al. (2014) find a decrease in child mortality of 25 percent. An intervention that allows beneficiaries to monitor health care providers even reduced child mortality by 33 percent (Björkman and Svensson, 2009). As outlined by Banerjee and Duflo (2011) and Dupas and Miguel (2017), there are highly effective and relatively cheap treatments for many diseases that affect the poor. For example, chlorine can largely reduce the probability to contract diarrhea. Oral rehydration solutions are a highly effective way to treat dehydration caused by diarrhea, a major cause of child mortality. Considering this context, the large effect sizes I observe in this study appear reasonable.

VII. Mechanisms

How exactly does bank presence affect health? To shed light on this question, I lay out a framework of potential mechanisms and then discuss which ones have bearing in the data. To motivate these mechanisms, I consider three determinants of poor health status in developing countries. First, households have limited resources to invest in health. For example, they are potentially not able to afford a nutritious diet or pay medical bills. As increased bank presence stimulates business activity and thereby income, households are potentially able to invest more. Second, households are limited in their ability to move available resources across time and states. For example, they might skip a necessary doctor's visit if they have not built up enough emergency savings, cannot take an emergency loan, or do not have health insurance. As bank presence increases, households might gain access to savings accounts, credit products, or health insurance.[6] Finally, households often have limited access to good health care. They could benefit from improved health care supply if banks provide credit constrained health care providers with access to loans. Previous research has demonstrated that bank presence can stimulate business activity and improve households' income (Burgess et al., 2005; Bruhn and Love, 2014). It has also examined the effect of access to savings and credit products for households (Dupas et al., 2018; Karlan and Zinman, 2010). This paper provides evidence on two largely understudied aspects of banking: households gain access to health insurance and health care providers gain access to credit.

[6] In the context of developing countries, banks often either directly offer health insurance or partner with local insurance companies to provide access to and verification of potential customers.

There is an important trade-off in this study. In contrast to previous work, the natural experiment allows me to realistically capture effects that emerge due to interactions between financial access by different type of agents, and due to a higher diffusion over a longer time frame. However, this means that I cannot exactly quantify the extent to which each mechanism contributed to improving health, which would require isolated exogenous variation in the respective mechanisms. For instance, the researcher would need exogenous variation in credit access for health care providers only. While providing isolated exogenous variation is out of this study's scope, it provides a promising avenue for future research. Figure 7 summarizes the mechanisms discussed in this paper. There could also be interactions between mechanisms. For instance, the increase in health care supply I observe in equilibrium (discussed below) is also consistent with health care providers adjusting supply to increased aggregated health care demand. I review these three potential mechanisms in the following sections.

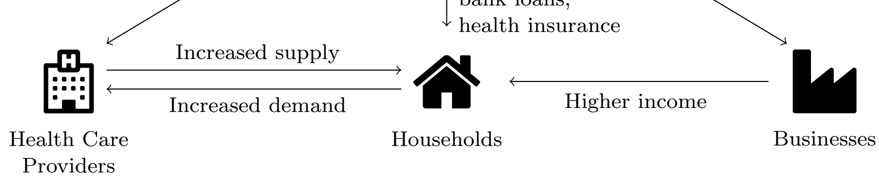


Figure 7. Mechanisms. This graphic depicts mechanisms driving the relationship between bank presence and health discussed in this paper. First, bank presence might stimulate business activity and thereby increase households' income; this could allow households to invest more in health. In particular, they might increase their health care demand. Second, bank presence might allow households to gain access to savings accounts, bank loans, or health insurance. This could allow households to allocate available resources across time and states. Finally, bank presence could increase credit access to health care providers, potentially improving supply. There could also be interactions between mechanisms, for instance, an increase in health care supply in equilibrium may be consistent with a response to higher aggregate health care demand.

A. Banks Stimulate Business Activity and Increase Households' Income

A first important determinant of poor health in developing countries is that households have limited resources to invest in health. For instance, they may be unable to afford a nutritious diet or pay medical bills. Previous literature has established that financial development in general and increased bank presence in particular can stimulate business activity and increase households' income (Bruhn and Love, 2014; Young, 2017),[7] which should in turn reflect in higher consumption.[8] Indeed, I observe that six years after the policy, households in treatment districts have 8 percent higher total consumption than those in control districts.[9] This finding is depicted in Column 1 of Table 6 and Figure 8(a). Next, I examine whether households increase their investments in health. Indeed, I observe that households spend 6 percent more on food, resulting in a quarter of a meal more per day (Table 6, Columns 2 and 3). This alone could have had positive impact on health. I find null effects on hygiene-related expenses, including the amount spent on soap, insecticides, or toilet articles (Table 6, Column 4). I also observe that outpatient expenses decrease and inpatient expenses remain stable (Table 6, Columns 5 and 6).

Based on these findings on outpatient and inpatient expenses, can one conclude there is a decrease in health care demand? No, because medical expenses are a bad proxy for historical health care demand. First, I do not measure historical medical expenses, but only a snapshot at the time of the survey, six years after the policy. By then, households may have already improved their health status, which would reflect in lower medical expenses. Second, prices could have decreased in equilibrium, but these are not observed in this study. Finally, increased access to health insurance could have decreased households' out-of-pocket share and thereby lowered medical expenses. Summarizing, it is likely that an income channel is at play; households have higher total consumption and spend more on food. Other outcomes of health status such as increased vaccination rates or increased

[7] Young (2017) demonstrates stimulated business activity for the 2005 RBI policy.

[8] I do not directly measure income, as this measure is often unreliable in survey data (Deaton and Zaidi, 2002).

[9] Note that this total increase in consumption is an equilibrium result. It not only reflects an increase in income through stimulated activity, but could also be influenced by households gaining access to financial instruments or being able to work more due to fewer illnesses.

Table 6: Households Have Higher Total and Food Consumption

	Total consumption (log Rs) (1)	Food consumption (log Rs) (2)	Meals per day (no.) (3)	Hygiene expenses (log Rs) (4)	Outpatient expenses (log Rs) (5)	Inpatient expenses (log Rs) (6)
Treated	0.07**	0.06*	0.24**	0.06	-0.45*	-0.14
	(0.04)	(0.03)	(0.10)	(0.06)	(0.23)	(0.30)
Control Mean	7.48	6.71	2.75	4.02	2.73	1.33
Change (%)	7.68	5.73	8.64	5.82	-36.06	-13.46
First Stage	0.75	0.71	0.68	0.66	0.70	0.56
Bandwidth	4,120	2,755	3,004	2,246	3,793	1,902
Efficient Obs.	14,903	11,415	16,611	9,896	17,418	8,537
Observations	21,410	21,345	34,773	23,010	29,182	27,312
Baseline Control	Yes	Yes	Yes	Yes	Yes	Yes

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. Data IHDS II (2011/2012). Household level. All variables in rupees are in per capita per month and trimmed at the 10th and 90th percentile. Hygiene expenses refer to amount spent on soap, insecticide, toilet articles, etc.

probability of deliveries at health care facilities point towards increased health care demand.

One note to make is that other studies that provide isolated exogenous variation in income find no effects on health. Haushofer and Shapiro (2018) conduct an RCT providing large unconditional cash transfers to Kenyan households. The average transfer size is 709 USD, around two years of per capita expenditure. Comparing recipients of cash transfers to non-recipients in the same village three years after the transfer, the authors find a 25 percent increase in consumption (47 USD). Note that this increase is more than three times as large as the consumption increase I observe in my natural experiment. Despite this strong increase in consumption, the authors do not detect an effect on health status in the short run (after nine months) or in the long run (after three years). Even if cash transfers are so high and distributed to so many households that they depict in total a fiscal shock of over 15 percent of the local GDP, researchers do not detect an impact on health status. Egger et al. (2019) find that cash transfers induced an increase in consumption by 13 percent after 18 months and identified other aggregate demand effects. However, the "estimated effects are close to zero and not significant for ... health." Of course, there are important differences between this study and unconditional cash-transfer RCTs, including the time elapsed before findings are measured. One cannot extrapolate from these two studies to the setting of this natural experiment, but they carefully suggest that other channels such as access to health insurance for households and access to credit for health care providers might play a role.

B. Households Gain Access to Financial Instruments

A second important determinant of poor health in developing countries is that households have limited ability to move available resources across time and states. For example, they might skip a necessary doctor's visit if they have not built up enough emergency savings, cannot take an emergency loan, or do not have health insurance. As bank presence increases, households might gain access to savings accounts, bank loans, or health insurance. As a first step, I utilize questions from the IHDS II to investigate whether households in treatment districts are more likely to take up these financial instruments due to the policy. Households are asked whether in the past five years they used any savings account, had any bank loan, or had any health insurance. Findings are reported in Table 7. Six years after the policy was introduced, households in treatment districts are 36 percent more likely to have used a savings account. In control areas, 51 percent of households have savings accounts. The discontinuity is depicted in Figure 8(b). Considering all households in the sample, I find a positive but insignificant effect on bank loans. This is in line with other studies on credit impact that find low take-up (Banerjee et al., 2015b). Importantly, households also gain access to health insurance; based on the low control mean of six percent, I indeed observe a large increase on this dimension (278%). The respective discontinuity is depicted in Figure 8(c). Take-up is balanced pre-policy as indicated in Table A22. Not all outcome variables are available pre-policy, in which case similar dimensions of financial access are shown to be smooth.

To examine whether households that took up financial instruments experienced a stronger effect on health, I utilize the pre-policy survey to predict the probability of take-up. Consistent with financial instruments having played an important role, households in the upper half of these probability distributions show stronger positive effects on health for all three instruments (Table A27). Coefficients for households in the lower half of the distribution—

Table 7: Households Gain Access to Savings Accounts and Health Insurance

	Savings account (yes/no) (1)	Bank loan (yes/no) (2)	Health insurance (yes/no) (3)
Treated	0.19*	0.04	0.17**
	(0.10)	(0.05)	(0.07)
Control Mean	0.51	0.23	0.06
Change (%)	36.48	19.70	272.69
First Stage	0.69	0.66	0.56
Bandwidth	3,023	2,370	1,704
Efficient Obs.	16,674	12,856	8,482
Observations	36,786	36,785	34,181
Baseline Control	No	No	No

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. Data IHDS II (2011/2012). Household level. Questions ask whether the household used or owned the respective instrument in the past five years.

those less likely to take up the products—are smaller, and they are significant for savings accounts and bank loans but insignificant for health insurance. This evidence particularly suggests that health insurance could have played a key role. Note that financial devices are not randomly distributed in these tests. This means that the predicted probability of take-up correlates with unobservables potentially driving the treatment strength. I thus benchmark my findings to studies that provide isolated exogenous variation in access to financial services. Conducting RCTs in three countries, Dupas et al. (2018) conclude that providing basic savings accounts alone is unlikely to improve welfare noticeably. Evidence on (micro-)credit products suggests that they are also unlikely to improve health (Angelucci et al., 2015; Augsburg et al., 2015; Attanasio et al., 2015; Banerjee et al., 2015a; Crépon et al., 2015; Karlan and Zinman, 2010; Tarozzi et al., 2015). Finally, studies on the impact of health insurance on health status are sparse, but partly suggest positive effects (Erlangga et al., 2019). This implies that especially access to health insurance could have played an important role in improving health.

C. Health Care Providers Gain Access to Credit

A third determinant of poor health in developing countries is that households often have limited access to good health care services. I thus investigate whether increased bank presence allowed health care providers to gain access to credit and whether health care supply increased in equilibrium. Given the context of the health care system in India, making progress on access to or quality of health care might have significant consequences for health status. 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 for them. Fifty-two percent report that personnel absenteeism is a big issue and 53 percent have large problems with drug availability at health care facilities. If bank presence allows health care providers to relax their credit constraints, this could allow investing into new health care facilities, providing monetary incentives for medical personnel to decrease absenteeism rates, or purchasing drugs on stock. In order for bank presence to trigger an increase in supply, two conditions need to be satisfied. The first is that health care providers generally utilize bank loans to finance themselves. To examine whether this condition holds true, I use two data sets: the Prowess database and the Economic Census. The Prowess database provides detailed financial information about a sample of relatively large health care providers. The Economic Census only contains crude measures of balance sheet data but lists all health care provider establishments.

The Prowess database lists 388 companies conducting hospital activities. These companies have a broad asset range of USD 2,000 to 410 million, with a median asset size of 3 million. The average company has a bank loan of USD 3.29 million, corresponding to 22 percent of its total assets. As expected, private companies rely more heavily on bank loans, averaging 25 percent of their assets, but government companies still report bank loans of 19 percent of their assets. While there is a positive relationship between the share of bank loans over assets and the size of a company, there are many smaller companies that have a relatively high share (see Figure A2). Additionally, a high dependency on bank loans is not unique to companies that conduct hospital activities. A similar picture emerges for the 22 companies in the data that offer medical or dental activities. These have a narrower asset range of USD 115,000 to 107 million and a similar median asset size of 3 million. The average company has a bank loan of USD 1.49 million, corresponding to 21 percent of its assets. Examining the Prowess data base allows the conclusion that relatively large companies heavily rely on bank financing.

To examine whether smaller health care 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 health care providers: only 1.59 percent of establishments with hospital activities and 2.00 percent of establishments with medical or dental activities list institutional loans as their major source of finance.[10] Instead, commonly cited major sources of finance for establishments with hospital activities (medical or dental activities) are self-finance with 44 percent (72 percent) and government sources with 39 percent (12 percent). That few health care providers cite institutional loans as their major source of finance does not imply that they do not rely on bank loans. Health care 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, just like regular businesses, they rely on bank loans.

The second condition that needs to be satisfied is that health care providers are credit constrained. Otherwise, they would either not take up the extra bank loans or only substitute more expensive credit. While there is no evidence available for health care 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. I conclude that the conditions are met for credit access to health care providers to play a role.

Turning towards the regression discontinuity analysis, I examine whether health care providers are more likely to cite institutional loans as their major source of finance in

[10] Note that institutional loans are likely to refer to bank loans here, since other major loan distributors such as money lenders are listed under a category of non-institutional loans.

Table 8: Hospitals Increasingly Use Institutional Loans

	Major source of finance				
	Institutional loan (share) (1)	Self-finance (share) (2)	Government sources (share) (3)	Donations (share) (4)	Non-institutional loan (share) (5)
Treated	0.010**	-0.058	0.053	-0.010	0.002
	(0.004)	(0.048)	(0.047)	(0.016)	(0.003)
Control Mean	0.014	0.495	0.332	0.085	0.002
Change (%)	67.77	-11.67	16.06	-11.42	87.01
First Stage	0.79	0.80	0.80	0.80	0.81
Bandwidth	2,435	4,393	3,928	4,018	5,044
Efficient Obs.	163	272	245	248	303
Observations	538	539	539	539	541
Baseline Control	No	No	No	No	No

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. Data Economic Census (2013). District level. An outcome is the share of hospitals in a given district that state their major source of finance is a certain source, e.g., institutional loans.

treatment than in control districts. Running a district-level regression, I find that eight years after the policy, the share of establishments with hospital activities that cite an institutional loan is significantly higher in treatment districts. The finding is depicted in Column 1 of Table 8, and the respective discontinuity is shown in Figure 8(d). The coefficient on establishments that conduct medical or dental activities is insignificant. Other forms of financing such as self-finance are not increasingly more likely to be reported after the policy. This can cautiously be interpreted against evidence of other drivers of supply growth, including an aggregate demand effect or that hospital owners generally become richer due to increased business activity and thus finance hospitals from their own pockets.

If health care providers are indeed credit constrained, I expect the increased take-up of institutional loans to translate into an improvement in health care supply. Note that I observe health care supply in the equilibrium and do not isolate the effect of increased credit access to health care providers on supply. I find that eight years after the policy was introduced, treatment districts have significantly more hospitals. As described in Column 3 of Table 9, treatment districts have 140 percent more hospitals than control districts, corresponding to 74 hospitals per 100,000 people versus the control mean of 31 hospitals per 100,000 people. While the control mean appears large, these hospitals are relatively small; they have on average only seven employees listed in the census. The

Table 9: Hospitals Open

	Pre-policy (2005)		Post-policy (2013)	
	Hospitals (log no.) (1)	Other medical service providers (log no.) (2)	Hospitals (log no.) (3)	Other medical service providers (log no.) (4)
Treated	-0.15	0.26	0.88***	0.10
	(0.16)	(0.31)	(0.33)	(0.35)
Control Mean	5.42	5.22	5.96	5.28
Change (%)	-13.96	29.96	140.07	10.55
First Stage	0.80	0.80	0.80	0.81
Bandwidth	4,328	3,176	3,127	3,417
Efficient Obs.	268	203	201	213
Observations	539	538	538	538
Baseline Control	No	No	No	No

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. Data Economic Census (2005 and 2013). District level. All variables are transformed into log form and winsorized at the 1st and 99th percentile.

respective discontinuity is depicted in Figure 8(e). The coefficient for establishments that conduct medical or dental activities is insignificant. These facilities are even smaller with an average two employees per establishment. The presence of both hospitals and other medical service providers is balanced on baseline as observable in Column 1 and 2.

I further complement this information from the Economic Census with household survey evidence from the DHS (2015/2016). I utilize questions that ask whether the household has big problems with health care facilities in terms of access (distance, transport) and quality (personnel absenteeism, lack of female health care personnel, lack of drugs at the facility). As becomes evident studying the control means in Table 10, a high share of households are unsatisfied with the health care system. Here I focus on the urban sample of the population; coefficients are mostly insignificant for the rural sample. Twenty percent of urban households see distance to a facility as a big problem and 17 percent evaluate the transport to a facility as very problematic. In terms of quality, 44 percent see absenteeism of personnel as a big issue. Thirty-seven percent find lack of specifically female personnel concerning. Finally, another large issue is related to availability of drugs; 45 percent of households evaluate this as a big problem. Importantly, ten years after the policy, households in urban areas of treatment districts are significantly less likely to report big problems with health care providers. Probabilities for access being an issue are

Table 10: Households Report Less Problems

	Big problem with health care providers				
	Access		Quality		
	Distance to facility (yes/no) (1)	Taking transport to facility (yes/no) (2)	No personnel at facility (yes/no) (3)	No female personnel at facility (yes/no) (4)	No drugs at facility (yes/no) (5)
Treated	-0.12***	-0.11***	-0.14**	-0.20**	-0.15**
	(0.04)	(0.04)	(0.06)	(0.08)	(0.07)
Control Mean	0.20	0.17	0.44	0.37	0.45
Change (%)	-57.66	-65.35	-32.39	-54.27	-32.35
First Stage	0.60	0.57	0.62	0.62	0.59
Bandwidth	2,053	1,922	2,216	2,258	2,015
Efficient Obs.	34,937	34,395	41,751	42,131	34,829
Observations	128,525	128,525	129,568	129,568	128,525
Baseline Control	No	No	No	No	No

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. Data DHS (2015/2016). Urban sample. Household level.

58 and 65 percent lower in treatment than control districts. In terms of concerns with quality, probabilities are 32 to 54 percent lower in treatment districts (see Figure 8(f) for the discontinuity of personnel absence). These findings supplement the evidence from the Economic Census, further suggesting that health care supply improved as a result of the policy.

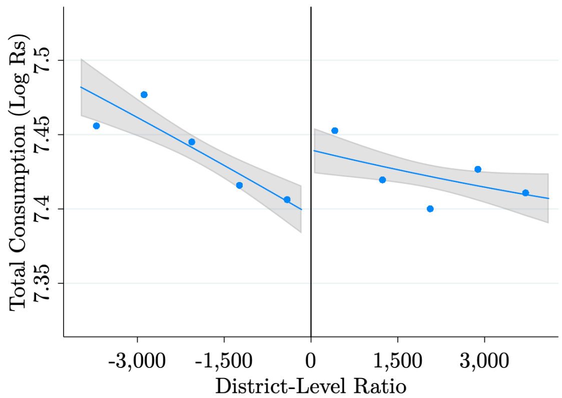
Finally, in alignment with the credit access narrative, I observe a stronger reaction for private hospitals on loan take-up and supply increase. Private hospitals account for approximately three quarters of all hospitals in the country. Findings are indicated in Table 11. Eight years after the policy, treatment districts have an 88 percent higher share of private hospitals that cite an institutional loan as their main source of finance than control districts. No government hospital cites an institutional loan as their main source of finance. However, from the Prowess data, I know that bank loans are still highly relevant for government hospitals, even though unlikely to be the major source of finance. Any effect on government hospitals is hidden due to the lack of more detailed balance sheet data in the census. Examining the response in supply, I find that treatment districts have 130 percent more private hospitals than control districts, but only 81 percent more government hospitals after the policy. Corresponding to the data of the Economic Census,

Table 11: Stronger Reaction for Private Hospitals

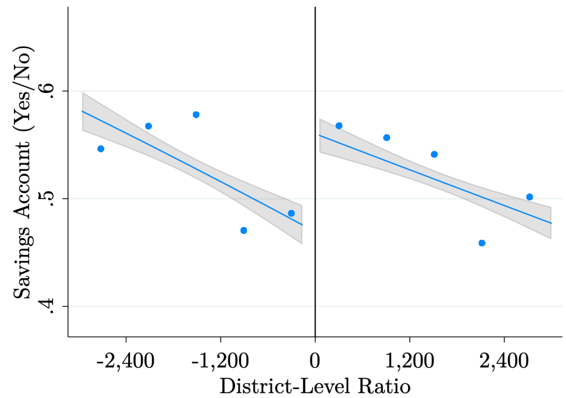
	Private		Government	
	Institutional loan (share) (1)	Hospitals (log no.) (2)	Institutional loan (share) (3)	Hospitals (log no.) (4)
Treated	0.020**	0.84**	-	0.64*
	(0.009)	(0.36)	(-)	(0.33)
Control Mean	0.025	5.27	0.000	4.41
Change (%)	87.52	16.02	-	14.63
First Stage	0.79	0.81	-	0.81
Bandwidth	2,357	3,382	-	3,633
Efficient Obs.	156	211	-	226
Observations	528	538	-	539
Baseline Control	No	No	-	No

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. Data Economic Census (2013). District level. Outcomes in Columns 1 and 3 are shares of private or government hospitals in a given district that state their major source of finance is a certain source, e.g., institutional loans. Variables in Columns 2 and 4 are transformed into log form and winsorized at the 1st and 99th percentile.

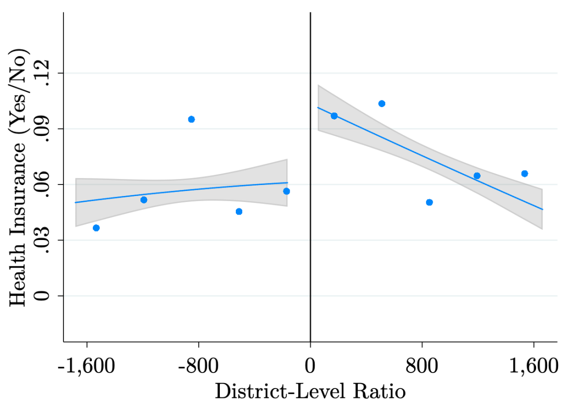
I find that households are significantly more likely to state that they generally go to private providers for treatment. Households partly substitute away from government providers. Evidence for this is depicted in Table A26 (DHS (2015/2016)). Summarizing, I observe a stronger credit take-up and supply-side reaction for private providers, the group that is more likely to benefit from access to bank loans. Finally, qualitative interviews with Indian bank employees support the hypothesis that banking services post 2005 allowed the health care sector to grow.



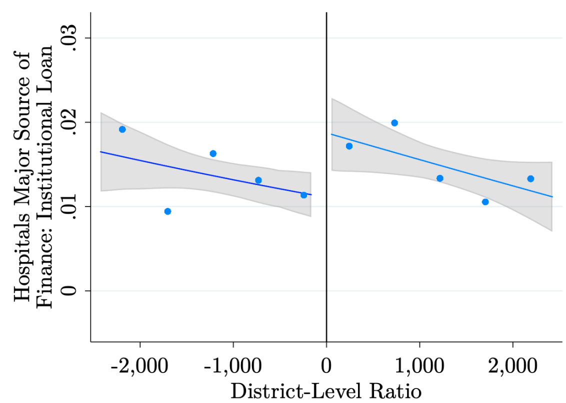
(a) Post-Policy (2011/2012)



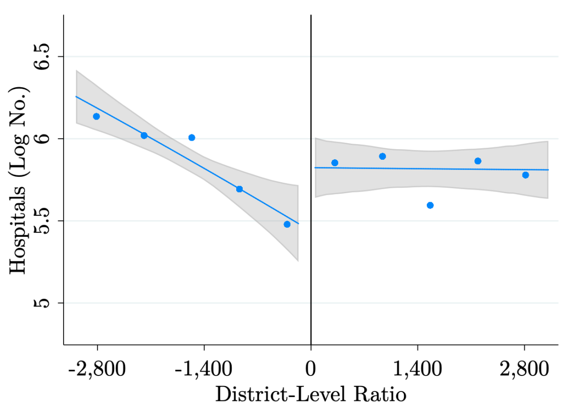
(b) Post-Policy (2011/2012)



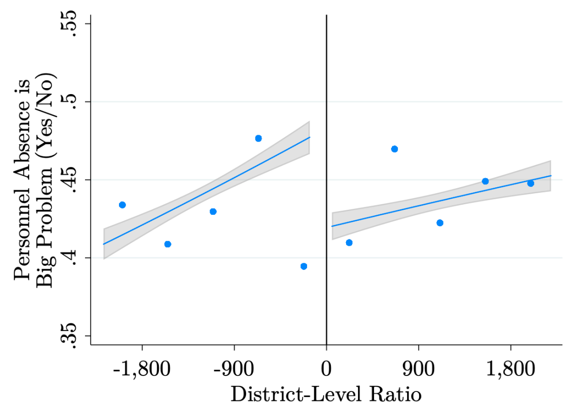
(c) Post-Policy (2011/2012)



(d) Post-Policy (2011/2012)



(e) Post-Policy (2011/2012)



(f) Post-Policy (2011/2012)

Figure 8. 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 90 percent confidence intervals indicated in gray.

VIII. Robustness and Placebo Tests

To demonstrate robustness of my results, I initially test whether coefficients remain statistically significant for different bandwidth choices. In a first approach, 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 for main outcomes are described in Table A10 and for mechanism outcomes in Table A11. Figure A4 provides a summary. For main outcomes, considering the optimal bandwidth ± 0.25, 75 to 88 percent of coefficients remain statistically significant. Examining bandwidth ± 0.50, 63 percent remain statistically significant. For mechanism outcomes, applying the bandwidth ± 0.25, 81 percent of coefficients remain statistically significant. Considering the optimal bandwidth ± 0.50, 56 percent remain 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 A12, A13, A14, and A15, I also consider an MSE-optimal selector that chooses separately bandwidths to the left and to the right of the cutoff. Additionally, I examine another selector suggested by Calonico et al. (2020), one the optimizes the coverage error rate (CER). Again, I examine the selector with identical and different bandwidths to the left and right of the cutoff. Figure A5 summarizes the results. For main outcomes, 69 to 88 percent of results remain statistically significant; for mechanism outcomes, 88 to 94 percent remain statistically significant. This suggests that results are robust to different bandwidth selectors.

Results are also robust taking into account possible bias corrections due to the MSE-optimal bandwidth selector, discussed by Calonico et al. (2014) and Cattaneo and Vazquez-Bare (2017). Results are depicted in Tables A12, A13, A14, as well as A15, and summarized in Figure A6. One hundred percent of the coefficients in both main and mechanism outcomes 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 either linear or quadratic approximations. Additionally, I examine robustness with respect to polynomials of degree three. Findings are described in Table A16 and A17 and summarized in Figure A7. For polynomials of degree two, 94 percent of main outcomes and 69 percent of mechanism outcomes remain statistically significant. For polynomials of degree three, not recommended by the current literature, I still find that 56 percent of main outcomes and 19 percent of mechanism outcomes remain significant. Summarizing, results are highly robust to alternative polynomials suggested by the econometric literature.

IX. Conclusion

In the preceding sections, I have shown first evidence that financial development in general and bank presence in particular can improve health. In contrast to previous RCTs that consistently find null effects on health, I observe effects that emerge due to interactions between financial access by different types of agents, and due to a higher diffusion over a longer time frame. In particular, households in treatment districts, which benefited from a policy of government incentives for banks to open new branches, have lower morbidity rates. This in turn positively impacts their economic situation; they miss fewer days of work or school due to an illness and spend significantly less on medical expenses. Ten years after the policy was introduced, I observe persistently lower morbidity rates, higher vaccination rates, and lower risks associated with pregnancies. Additionally, I provide evidence that two previously understudied aspects of banking played a role: households gain access to health insurance and health care providers gain access to credit.

This paper has important implications for policy as well as for future research. Policymakers can conclude that it can be beneficial for the health of their citizens to incentivize banks to enter underserved locations. They might shift focus on the interaction of banks with local providers of services that policymakers want to foster. Indeed, the RBI announced a new policy in May 2021 to incentivize banks to quickly deliver credit to health care providers in light of the COVID crisis, announcing plans to inject USD 6.78 billion of liquidity. Beyond health, other services to consider are, for instance, education; just as health care providers, providers of education services are likely to be constrained. This paper also speaks to researchers, suggesting promising new areas of interest. One open question is to what extent the different mechanisms contribute to improving health, requiring exogenous variation in, for instance, credit access to health care providers only. A second line of inquiry is asking whether other dimensions of wellbeing such as education can be positively impacted by bank presence. Gaining an understanding of these questions could significantly advance our knowledge of the impact of bank presence and the scope of policymakers to improve their citizens' wellbeing.

REFERENCES

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

Asher, S., Lunt, T., Matsuura, R., and Novosad, P. (2021). Development Research at High Geographic Resolution: An Analysis of Night Lights, Firms, and Poverty in India using the SHRUG Open Data Platform. World Bank Group Policy Research Working Paper 9540.

Attanasio, O., Augsburg, B., De Haas, R., Fitzsimons, E., and Harmgart, H. (2015). The impacts of microfinance: Evidence from joint-liability lending in Mongolia. *American Economic Journal: Applied Economics*, 7(1):90–122.

Augsburg, B., De Haas, R., Harmgart, H., and Meghir, C. (2015). The impacts of microcredit: Evidence from Bosnia and Herzegovina. *American Economic Journal: Applied Economics*, 7(1):183–203.

Banerjee, A. and Duflo, E. (2011). *Poor Economics*. New York.

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

Banerjee, A., Karlan, D., and Zinman, J. (2015b). Six randomized evaluations of microcredit: Introduction and further steps. *American Economic Journal: Applied Economics*, 7(1):1–21.

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

Beck, T., Demirgüç-Kunt, A., and Maksimovic, V. (2005). Financial and legal constraints to growth: does firm size matter? *The Journal of Finance*, 60(1):137–177.

Bharadwaj, P., Jack, W., and Suri, T. (2019). Fintech and household resilience to shocks: Evidence from digital loans in Kenya. National Bureau of Economic Research Working Paper 25604.

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

Björkman-Nykvist, M., Andrea, G., Svensson, J., and Yanagizawa-Drott, D. (2014). Evaluating the impact of the living goods entrepreneurial model of community health delivery in Uganda: A cluster-randomized controlled trial. Technical report, Mimeo.

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

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

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

Burgess, R., Pande, R., and Wong, G. (2005). Banking for the poor: Evidence from India. *Journal of the European Economic Association*, 3(2-3):268–278.

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

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

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

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

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

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

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

Deaton, A. and Zaidi, S. (2002). Guidelines for constructing consumption aggregates for welfare analysis. World Bank LSMS Working Paper 135.

Dupas, P., Karlan, D., Robinson, J., and Ubfal, D. (2018). Banking the unbanked? evidence from three countries. *American Economic Journal: Applied Economics*, 10(2):257–97.

Dupas, P. and Miguel, E. (2017). Impacts and determinants of health levels in low-income countries. In Banerjee, A. V. and Duflo, E., editors, *Handbook of Economic Field Experiments*, volume 2, pages 3–93. Elsevier, Amsterdam.

Dupas, P. and Robinson, J. (2013). Why don't the poor save more? evidence from health savings experiments. *American Economic Review*, 103(4):1138–71.

Egger, D., Haushofer, J., Miguel, E., Niehaus, P., and Walker, M. W. (2019). General equilibrium effects of cash transfers: experimental evidence from Kenya. National Bureau of Economic Research Working Paper 26600.

Erlangga, D., Suhrcke, M., Ali, S., and Bloor, K. (2019). The impact of public health insurance on health care utilisation, financial protection and health status in low-and middle-income countries: A systematic review. *PloS one*, 14(8):e0219731.

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

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

Goldsmith, R. W. (1969). *Financial Structure and Development.* Yale University Press, New Haven.

Haushofer, J. and Shapiro, J. (2013). Household response to income changes: Evidence from an unconditional cash transfer program in Kenya. Working paper.

Haushofer, J. and Shapiro, J. (2018). The long-term impact of unconditional cash transfers: experimental evidence from Kenya. Working Paper.

Henderson, J. V., Storeygard, A., and Weil, D. N. (2011). A Bright Idea for Measuring Economic Growth. *American Economic Review.*

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

King, R. G. and Levine, R. (1993). Finance and growth: Schumpeter might be right. *The Quarterly Journal of Economics*, 108(3):717–737.

Kremer, M., Leino, J., Miguel, E., and Zwane, A. P. (2011). Spring cleaning: Rural water impacts, valuation, and property rights institutions. *The Quarterly Journal of Economics*, 126(1):145–205.

Levine, R. (2005). Finance and growth: theory and evidence. In Aghion, P. and Durlauf, S. N., editors, *Handbook of Economic Growth*, volume 1A, pages 865–934. Elsevier, Amsterdam.

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

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

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

Rajan, R. and Zingales, L. (1996). Financial dependence and growth. National Bureau of Economic Research Working Paper 5758.

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

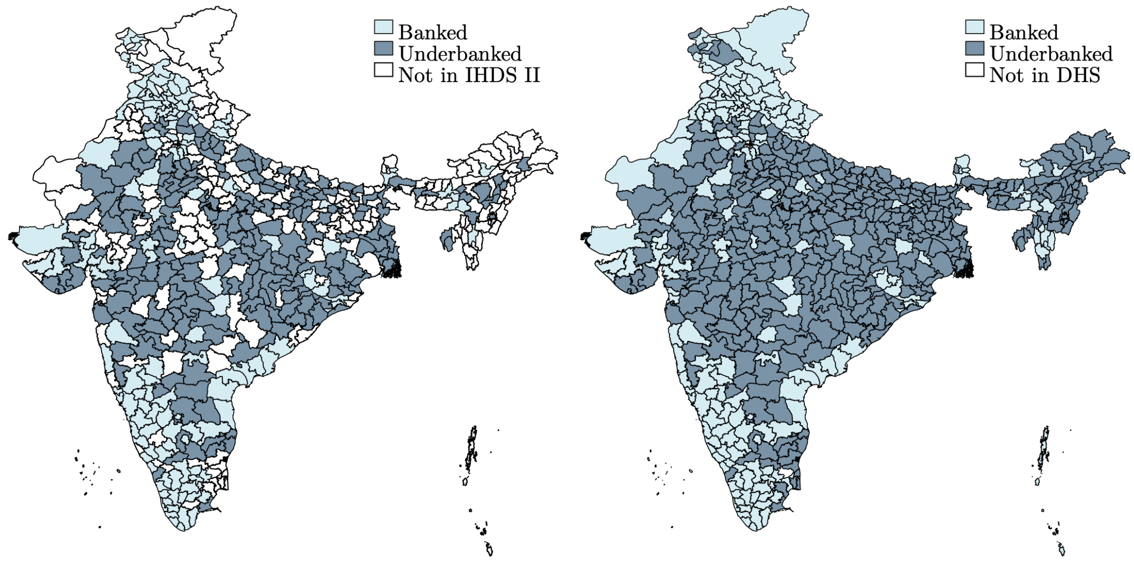
Young, N. (2017). Banking and growth: Evidence from a regression discontinuity analysis. EBRD Working Paper 207.

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)	(140)	(59)	(67)	(102)	(120)
Branches (no.)	65	71	103	172	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. Between 2004, one year before the policy, and 2016, the final year of the last survey, I generally observe a large branch growth of 142 percent in the average district. Districts with a population-to-branch ratio in the range of $\pm 3,000$ of the policy cutoff have a slightly higher number of branches on average.

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.



(a) IHDS II

(b) DHS

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

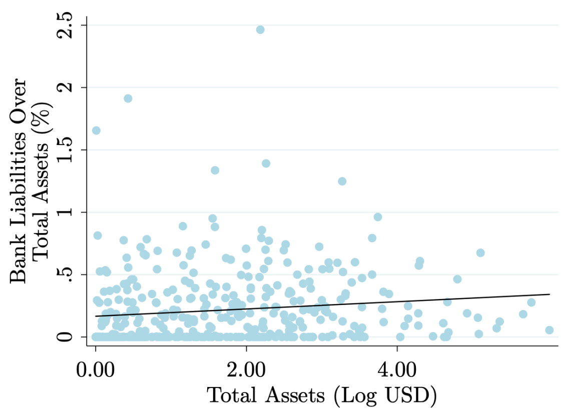


Figure A2. Relationship between Bank Liabilities and Total Assets. As expected, there is a positive relationship between the share of bank liabilities over total assets and the size of the company proxied by total assets. However, there are many companies of lower asset size that have a relatively high share of bank liabilities over total assets.

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)	705	699	1,841	1,828
	(315)	(309)	(823)	(804)
Food consumption (Rs)	372	370	848	844
	(121)	(119)	(281)	(278)
Financial Access				
Savings account (yes/no)			0.57	0.53
			(0.49)	(0.50)
Any loan (yes/no)	0.46	0.48	0.60	0.62
	(0.50)	(0.50)	(0.49)	(0.49)
Any bank loan (yes/no)			0.22	0.23
			(0.41)	(0.42)
Largest loan from bank (yes/no)	0.11	0.12	0.18	0.17
	(0.32)	(0.32)	(0.38)	(0.38)
Largest loan amt (Rs)	4,482	4,862	15,448	17,134
	(8,698)	(9,048)	(25,365)	(26,498)
Health insurance (yes/no)	0.03	0.02	0.11	0.08
	(0.16)	(0.15)	(0.31)	(0.28)
Health				
Days ill (yes/no)	0.47	0.45	0.55	0.52
	(0.50)	(0.50)	(0.50)	(0.50)
Days ill (no.)	2.78	2.54	3.23	2.97
	(4.08)	(3.89)	(4.19)	(4.00)
Days missed (yes/no)	0.36	0.34	0.40	0.39
	(0.48)	(0.47)	(0.49)	(0.49)
Days missed (no.)	1.46	1.36	1.64	1.60
	(2.62)	(2.51)	(2.76)	(2.73)
Treatment spending (yes/no)	0.45	0.43	0.53	0.51
	(0.50)	(0.49)	(0.50)	(0.50)
Treatment spending (Rs)	43	41	126	121
	(82)	(80)	(204)	(202)
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 trimmed at the 10th and 90th 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. Generally speaking, I observe a positive trend in consumption measures and financial access, while health status remained stable over the period between 2004/2005 and 2011/2012. Assume that general consumption measures are positively correlated with being more sensitive or informed about illnesses. This would explain that health status does not improve, e.g., the number of days ill would have an upward bias due to self-reporting. Notice that this bias would go in the other direction than the effects I detect; in treatment districts, consumption measures show increases, but health measures show decreases. Self-reporting effects thus make it potentially less likely for me to find an effect on health. I complement evidence on health status with measures that are not self-reported. Additionally, I observe that households in districts within the range of -3,000 to +3,000 of the normalized ratio are remarkably similar to households in all districts, strengthening external validity of my design.

Table A3: Households Summary Statistics (DHS)

	DHS 2015/2016	
	All districts	[-3,000,+3,000]
	(1)	(2)
Morbidity		
Sick child (yes/no)	0.27	0.26
	(0.45)	(0.44)
Health Care Visits		
Any reason (yes/no)	0.28	0.28
	(0.45)	(0.45)
Children's treatment (yes/no)	0.11	0.10
	(0.31)	(0.30)
Women's treatment (yes/no)	0.16	0.16
	(0.37)	(0.37)
Facility delivery (yes/no)	0.02	0.02
	(0.13)	(0.13)
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)
Pregnancies		
Experienced miscarriage (yes/no)	0.04	0.04
	(0.20)	(0.19)
Experienced stillbirth (yes/no)	0.01	0.00
	(0.08)	(0.07)
Health Care 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	576	576

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.

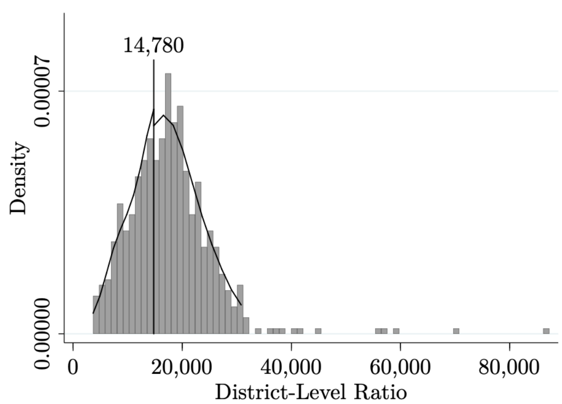


Figure A3. McCrary (2008) Density Test. There is no evidence of manipulation around the cutoff. The McCrary estimator is -0.1998 with a p-value of 0.8416; I do not reject smoothness around the cutoff.

Table A5: **Smoothness Pre-Policy: Economic Activity and Population Characteristics**

	1990	1991	...	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005
Nightlights															
Total light (log)				-0.07	-0.03	0.03	0.06	-0.03	0.16	0.02	0.05	-0.16	-0.00	-0.13	-0.06
				(0.25)	(0.27)	(0.28)	(0.28)	(0.28)	(0.27)	(0.28)	(0.31)	(0.27)	(0.30)	(0.29)	(0.29)
Economic Census															
Empl. (log no.)	-0.16							-0.04							0.07
	(0.25)							(0.15)							(0.13)
Empl. manuf. (log no.)	-0.05							-0.04							0.02
	(0.19)							(0.14)							(0.16)
Empl. services (log no.)	-0.16							0.03							0.06
	(0.24)							(0.11)							(0.13)
Population Census															
Pop. (log no.)		0.01									-0.00				
		(0.11)									(0.10)				
Pop. rural (log no.)		0.01									0.00				
		(0.10)									(0.10)				
Pop. urban (log no.)		-0.11									-0.06				
		(0.08)									(0.08)				
Pop. literate (log no.)		-0.05									-0.07				
		(0.14)									(0.11)				
Tar road (yes/no)		-0.08									0.04				
		(0.07)									(0.06)				

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. 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. Population measures the total population, followed by a split in rural and urban. The last two variables from the population census measure the 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
Change (%)	284.06	46.22	90.26
First Stage	0.54	0.66	0.61
Bandwidth	1,633	2,363	1,982
Efficient Obs.	8,104	12,862	9,783
Observations	34,415	36,805	34,832
Baseline Control	Yes	Yes	Yes

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. Data IHDS II (2011/2012). Household level.

Discussion A2. The Ministry of Health and Family Welfare is a government agency that implements policies related to health. 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 supply of health care 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 pose a threat to identification. However, in 2013 the ministry published a list of 184 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 difference in percent of priority districts in treatment and control districts within the bandwidth is seven percentage points. The correlation coefficient within the bandwidth is low at 0.08.

Another ministry that introduced health-related policies is the Ministry of Women and Child Development. Two policies in particular are worth considering in this context: the Integrated Child Development Services (ICDS) program and the ICDS Systems Strengthening and Nutrition Improvement Project (ISSNIP). The ICDS was introduced in 1975 and has, among other goals, the objective to reduce 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 200 priority districts. In 2013/2014, a second wave of rollout followed in another 200 districts. Priority districts were defined based on nutritional status of children and anemia level among pregnant women. Unfortunately, only the list of the 200 districts in the first wave is available. In 2012, around the same time that the strengthened ICDS was rolled out, the ministry implemented the ISSNIP. This policy had the objective to shift the focus of the ICDS scheme to younger children. It focuses on 162 priority districts, also defined on the basis of undernutrition measures. Both policies have negative and insignificant coefficients, meaning that they were not significantly more likely to be implemented in treatment districts. Additionally, the difference in percent of priority districts in treatment and control districts within the bandwidth is 5 percentage points for ICDS wave one and 11 percentage points for ISSNIP. Correlation coefficients within the bandwidth are low at 0.06 and 0.13 respectively.

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 coefficients that are negative and statistically insignificant. The difference in percent of priority districts in treatment and control districts within the bandwidth is 23 for the first wave and -6 for the second wave. Correlation coefficients within the bandwidth are 0.25 and -0.07 respectively.

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

	Priority districts				
	NHM (yes/no) (1)	ICDS (1st wave) (yes/no) (2)	ISSNIP (yes/no) (3)	NREGA (1st wave) (yes/no) (4)	NREGA (2nd wave) (yes/no) (5)
Treated	0.21	-0.14	-0.23	-0.25	-0.02
	(0.20)	(0.19)	(0.19)	(0.23)	(0.25)
Control Mean	0.18	0.25	0.15	0.16	0.24
Change (%)	118.66	-57.84	-152.46	-151.04	-8.59
First Stage	0.70	0.77	0.78	0.70	0.67
Bandwidth	2,671	4,160	4,595	2,706	2,290
Efficient Obs.	176	260	290	181	151
Observations	581	581	581	581	581
Baseline Control	No	No	No	No	No

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. Data Ministry of Health and Family Welfare, Ministry of Women and Child Development, Ministry of Rural Development. District level. Regressions do not include state-level fixed effects.

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

	Priority districts				
	NHM (1)	ICDS (1st wave) (2)	ISSNIP (3)	NREGA (1st wave) (4)	NREGA (1st wave) (5)
All sample					
Total priority districts (no.)	169	180	156	196	125
Total priority districts (%)	29	31	27	34	22
Priority districts above cutoff (no.)	135	142	136	170	85
Priority districts above cutoff (%)	36	38	36	45	23
Priority districts below cutoff (no.)	34	38	20	26	40
Priority districts below cutoff (%)	17	19	10	13	20
Corr priority district and 1[Above]	0.20	0.20	0.28	0.33	0.04
Within bandwidth [-4,000,+4,000]					
Total priority districts (no.)	58	68	57	71	47
Total priority districts (%)	23	27	23	28	19
Priority districts above cutoff (no.)	37	42	39	54	23
Priority districts above cutoff (%)	26	29	27	38	16
Priority districts below cutoff (no.)	21	26	18	17	24
Priority districts below cutoff (%)	19	24	16	15	22
Corr priority district and 1[Above]	0.08	0.06	0.13	0.25	-0.07

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. Data Ministry of Health and Family Welfare, Ministry of Women and Child Development, Ministry of Rural Development. 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: Placebo Test: Regional Rural Banks

	Post-Policy (2010)	
	Branch Licenses (log no.) (1)	Branches (log no.) (2)
Treated	-0.54	-0.08
	(0.48)	(0.48)
Control Mean	1.39	1.05
Change (%)	-41.94	-7.63
First Stage	0.80	0.80
Bandwidth	2,812	2,959
Efficient Obs.	187	195
Observations	561	561
Baseline Control	Yes	Yes

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

Table A10: Robustness to Different Bandwidth Multipliers: Main Results

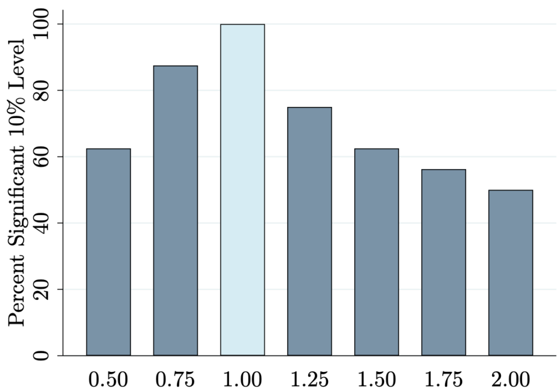
	Bandwidth Multiplier						
	x0.50	**x0.75**	**x1.00**	**x1.25**	**x1.50**	**x1.75**	**x2.00**
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Banks (Table 2)							
Branch licenses 2010 (log no.)	0.17**	0.21***	0.19***	0.17***	0.15***	0.14***	0.14***
	(0.08)	(0.06)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)
Branches 2010 (log no.)	0.15*	0.19***	0.17***	0.14**	0.13**	0.12**	0.14***
	(0.08)	(0.06)	(0.06)	(0.06)	(0.05)	(0.05)	(0.05)
Households health (Table 3)							
Days ill (yes/no)	-0.48	-0.26	-0.19**	-0.18***	-0.16**	-0.15**	-0.12**
	(0.35)	(0.17)	(0.09)	(0.07)	(0.07)	(0.06)	(0.06)
Days ill (log no.)	-0.55	-0.32*	-0.29**	-0.25**	-0.20*	-0.16	-0.13
	(0.41)	(0.17)	(0.12)	(0.12)	(0.11)	(0.11)	(0.10)
Days missed (yes/no)	-0.69	-0.39**	-0.30***	-0.28***	-0.25***	-0.22***	-0.19***
	(0.42)	(0.18)	(0.10)	(0.08)	(0.08)	(0.07)	(0.06)
Days missed (log no.)	-0.80*	-0.52***	-0.44***	-0.39***	-0.34***	-0.28***	-0.24***
	(0.44)	(0.20)	(0.13)	(0.12)	(0.11)	(0.10)	(0.09)
Medical expenses (yes/no)	-0.39	-0.21*	-0.18**	-0.17**	-0.15**	-0.13**	-0.11*
	(0.28)	(0.13)	(0.08)	(0.07)	(0.06)	(0.06)	(0.06)
Medical expenses (log Rs)	-1.65	-1.03**	-0.88**	-0.72**	-0.58*	-0.46*	-0.38
	(1.03)	(0.45)	(0.35)	(0.32)	(0.29)	(0.27)	(0.25)
Households health (Table 4)							
Vaccinated child (yes/no)	0.13***	0.11**	0.07*	0.05	0.04	0.02	0.02
	(0.05)	(0.05)	(0.04)	(0.03)	(0.03)	(0.03)	(0.03)
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)
HC visit (any reason) (yes/no)	-0.17***	-0.11***	-0.08**	-0.05*	-0.03	-0.02	-0.01
	(0.06)	(0.04)	(0.03)	(0.03)	(0.03)	(0.02)	(0.02)
HC visit (child's treatment) (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)
Households health (Table 5)							
Health care facility delivery (yes/no)	-0.001	0.003	0.005***	0.005***	0.005***	0.005***	0.004**
	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Experienced miscarriage (yes/no)	-0.020*	-0.013*	-0.010*	-0.010*	-0.009**	-0.008**	-0.007**
	(0.010)	(0.007)	(0.006)	(0.005)	(0.005)	(0.004)	(0.004)
Experienced stillbirth (yes/no)	-0.006**	-0.003**	-0.002*	-0.002*	-0.001	-0.001	-0.001
	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
HC visit (woman's treatment) (yes/no)	-0.109**	-0.074**	-0.051*	-0.036	-0.020	-0.007	-0.004
	(0.048)	(0.033)	(0.027)	(0.024)	(0.022)	(0.020)	(0.018)

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. For details of the regression refer to the respective main table. Summarized in Figure A4.

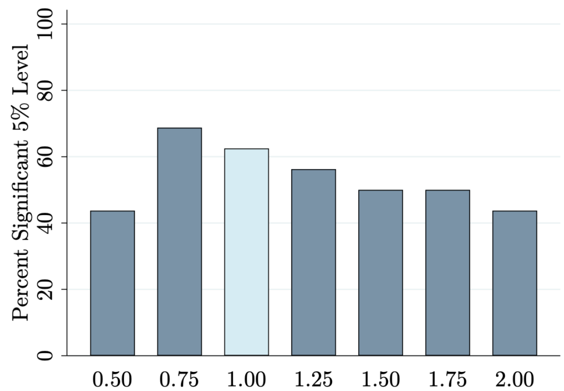
Table A11: **Robustness to Different Bandwidth Multipliers: Mechanisms Results**

	Bandwidth Multiplier						
	x0.50	**x0.75**	**x1.00**	**x1.25**	**x1.50**	**x1.75**	**x2.00**
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Households consumption (Table 6)							
Total consumption (log Rs)	0.10**	0.08**	0.07**	0.05	0.04	0.03	0.03
	(0.04)	(0.04)	(0.04)	(0.03)	(0.03)	(0.03)	(0.03)
Food consumption (log Rs)	-0.00	0.04	0.06*	0.06*	0.06*	0.05	0.04
	(0.06)	(0.04)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Meals per day (no.)	0.42*	0.25**	0.24**	0.25***	0.26***	0.24***	0.23***
	(0.23)	(0.12)	(0.10)	(0.09)	(0.09)	(0.08)	(0.07)
Outpatient expenses (log Rs)	-0.43	-0.58**	-0.45*	-0.35*	-0.25	-0.21	-0.16
	(0.40)	(0.26)	(0.23)	(0.21)	(0.19)	(0.16)	(0.14)
Households financial access (Table 7)							
Savings account (yes/no)	0.21	0.24*	0.19*	0.17**	0.12*	0.08	0.06
	(0.22)	(0.13)	(0.10)	(0.08)	(0.07)	(0.07)	(0.06)
Health insurance (yes/no)	0.25	0.25**	0.17**	0.12**	0.08*	0.05	0.04
	(0.18)	(0.12)	(0.07)	(0.05)	(0.04)	(0.04)	(0.03)
Health care supply (Table 8, 9, and 11)							
Institutional loan (share)	0.008	0.008	0.010**	0.010**	0.011**	0.011***	0.011***
	(0.006)	(0.005)	(0.005)	(0.004)	(0.004)	(0.004)	(0.004)
Institutional loan (share) - private	0.002	0.015	0.020**	0.022***	0.020**	0.019**	0.019**
	(0.010)	(0.009)	(0.009)	(0.008)	(0.008)	(0.008)	(0.008)
Hospitals (log no.)	1.74**	1.27***	0.88***	0.62**	0.42*	0.29	0.21
	(0.77)	(0.45)	(0.33)	(0.27)	(0.24)	(0.22)	(0.20)
Hospitals (log no.) - private	2.01**	1.22**	0.84**	0.57*	0.34	0.23	0.15
	(0.87)	(0.47)	(0.36)	(0.31)	(0.27)	(0.25)	(0.22)
Hospitals (log no.) - government	1.38*	0.88**	0.64*	0.50*	0.41*	0.33	0.25
	(0.72)	(0.43)	(0.33)	(0.28)	(0.24)	(0.21)	(0.19)
Survey on problems (Table 10)							
Distance to facility (yes/no)	-0.33**	-0.21***	-0.12***	-0.04	-0.01	0.00	0.00
	(0.14)	(0.07)	(0.04)	(0.04)	(0.03)	(0.03)	(0.03)
Taking transport to facility (yes/no)	-0.37**	-0.19***	-0.11***	-0.04	-0.00	0.01	0.02
	(0.18)	(0.06)	(0.04)	(0.03)	(0.03)	(0.03)	(0.03)
No personnel at facility (yes/no)	-0.20*	-0.23***	-0.14**	-0.11*	-0.08	-0.06	-0.05
	(0.12)	(0.08)	(0.06)	(0.05)	(0.05)	(0.05)	(0.05)
No female personnel at facility (yes/no)	-0.34*	-0.31***	-0.20**	-0.15**	-0.11**	-0.09*	-0.08*
	(0.18)	(0.11)	(0.08)	(0.06)	(0.06)	(0.05)	(0.05)
No drugs at facility (yes/no)	-0.22	-0.19**	-0.15**	-0.10*	-0.06	-0.04	-0.03
	(0.14)	(0.08)	(0.07)	(0.06)	(0.05)	(0.05)	(0.05)

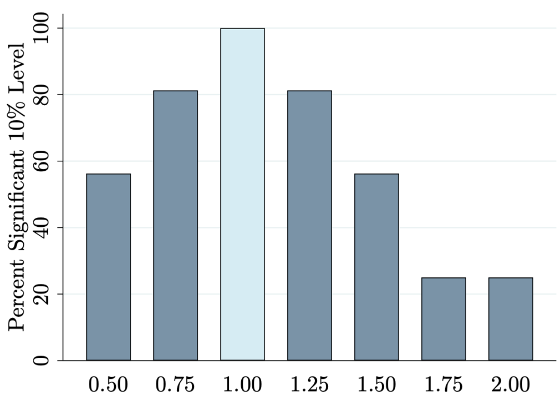
* p < 0.1, ** p < 0.05, *** p < 0.01. Standard errors in parentheses. For details of the regression refer to the respective main table. Summarized in Figure A4.



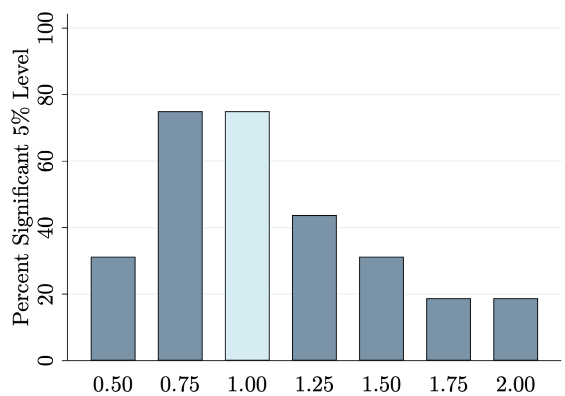
(a) Main Results



(b) Main Results



(c) Mechanisms Results



(d) Mechanisms Results

Figure A4. Percent of Results 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 A10 and A11.

Table A12: Robustness to Different Bandwidth Selectors: Main Results (1/2)

	MSE-optimal		CER-optimal	
	Common	Two-sided	Common	Two-sided
	(1)	(2)	(3)	(4)
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)
Households health (Table 3)				
Days ill (yes/no)	-0.19**	-0.13*	-0.26	-0.17
	(0.09)	(0.08)	(0.17)	(0.15)
	-0.21**	-0.16**	-0.28*	-0.19
	(0.09)	(0.08)	(0.17)	(0.15)
	-0.21**	-0.16	-0.28	-0.19
	(0.11)	(0.10)	(0.19)	(0.17)
Days ill (log no.)	-0.29**	-0.24*	-0.32*	-0.41
	(0.12)	(0.13)	(0.17)	(0.26)
	-0.36***	-0.29**	-0.35**	-0.45*
	(0.12)	(0.13)	(0.17)	(0.26)
	-0.36**	-0.29*	-0.35*	-0.45
	(0.15)	(0.17)	(0.19)	(0.29)
Days missed (yes/no)	-0.30***	-0.25***	-0.40**	-0.38**
	(0.10)	(0.09)	(0.19)	(0.17)
	-0.35***	-0.32***	-0.44**	-0.43**
	(0.10)	(0.09)	(0.19)	(0.17)
	-0.35***	-0.32***	-0.44**	-0.43**
	(0.12)	(0.11)	(0.21)	(0.20)
Days missed (log no.)	-0.44***	-0.42***	-0.52***	-0.57**
	(0.13)	(0.13)	(0.20)	(0.25)
	-0.54***	-0.51***	-0.59***	-0.64**
	(0.13)	(0.13)	(0.20)	(0.25)
	-0.54***	-0.51***	-0.59***	-0.64**
	(0.16)	(0.17)	(0.22)	(0.28)
Medical expenses (yes/no)	-0.18**	-0.14*	-0.21*	-0.19
	(0.08)	(0.07)	(0.13)	(0.14)
	-0.20***	-0.16**	-0.23*	-0.21
	(0.08)	(0.07)	(0.13)	(0.14)
	-0.20**	-0.16*	-0.23	-0.21
	(0.09)	(0.09)	(0.14)	(0.15)
Medical expenses (log Rs)	-0.88**	-0.63*	-1.04**	-0.94*
	(0.35)	(0.34)	(0.46)	(0.53)
	-1.10***	-0.83**	-1.19***	-1.10**
	(0.35)	(0.34)	(0.46)	(0.53)
	-1.10***	-0.83**	-1.19**	-1.10*
	(0.41)	(0.42)	(0.50)	(0.60)

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. 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 indicates 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 that is the bias-corrected RD estimator with 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 A5 and A6.

Table A13: **Robustness to Different Bandwidth Selectors: Main Results (2/2)**

	MSE-optimal		CER-optimal	
	Common (1)	Two-sided (2)	Common (3)	Two-sided (4)
Households health (Table 4)				
Vaccinated child (yes/no)	0.07*	0.06	0.11**	0.07
	(0.04)	(0.04)	(0.05)	(0.05)
	0.10**	0.07**	0.13**	0.08*
	(0.04)	(0.04)	(0.05)	(0.05)
	0.10**	0.07*	0.13**	0.08
	(0.05)	(0.04)	(0.06)	(0.05)
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)
HC visit (any reason) (yes/no)	-0.08**	-0.04	-0.12***	-0.09**
	(0.03)	(0.03)	(0.04)	(0.03)
	-0.11***	-0.07**	-0.14***	-0.11***
	(0.03)	(0.03)	(0.04)	(0.03)
	-0.11***	-0.07*	-0.14***	-0.11***
	(0.04)	(0.04)	(0.05)	(0.04)
HC visit (child's treatment) (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)
Households health (Table 5)				
Health care facility delivery (yes/no)	0.005***	0.006***	0.003	0.006***
	(0.002)	(0.002)	(0.002)	(0.002)
	0.006***	0.007***	0.003*	0.006***
	(0.002)	(0.002)	(0.002)	(0.002)
	0.006**	0.007***	0.003	0.006***
	(0.002)	(0.002)	(0.002)	(0.002)
Experienced miscarriage (yes/no)	-0.010*	-0.009*	-0.013*	-0.011*
	(0.006)	(0.005)	(0.007)	(0.006)
	-0.012**	-0.011**	-0.015**	-0.013*
	(0.006)	(0.005)	(0.007)	(0.006)
	-0.012*	-0.011*	-0.015*	-0.013*
	(0.007)	(0.007)	(0.008)	(0.007)
Experienced stillbirth (yes/no)	-0.002*	-0.001	-0.004**	-0.002*
	(0.001)	(0.001)	(0.001)	(0.001)
	-0.003**	-0.001	-0.004***	-0.003**
	(0.001)	(0.001)	(0.001)	(0.001)
	-0.003*	-0.001	-0.004**	-0.003*
	(0.001)	(0.001)	(0.002)	(0.002)
HC visit (woman's treatment) (yes/no)	-0.051*	-0.027	-0.076**	-0.060**
	(0.027)	(0.024)	(0.034)	(0.029)
	-0.077***	-0.051**	-0.094***	-0.078***
	(0.027)	(0.024)	(0.034)	(0.029)
	-0.077**	-0.051*	-0.094**	-0.078**
	(0.033)	(0.030)	(0.039)	(0.034)

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. 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 indicates 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 that is the bias-corrected RD estimator with 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 A5 and A6.

Table A14: Robustness to Different Bandwidth Selectors: Mechanism Results (1/2)

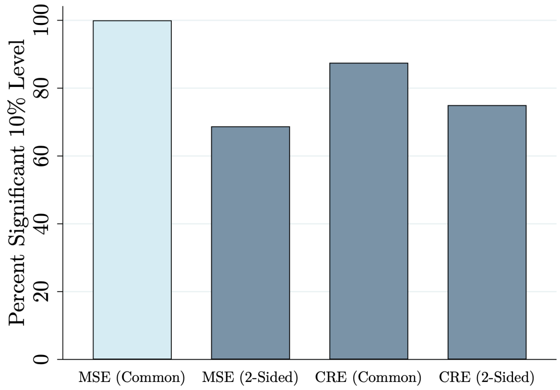
	MSE-optimal		CER-optimal	
	Common	Two-sided	Common	Two-sided
	(1)	(2)	(3)	(4)
Households consumption (Table 6)				
Total consumption (log Rs)	0.07**	0.07**	0.08**	0.10**
	(0.04)	(0.03)	(0.04)	(0.04)
	0.10***	0.10***	0.10***	0.12***
	(0.04)	(0.03)	(0.04)	(0.04)
	0.10**	0.10**	0.10**	0.12***
	(0.04)	(0.04)	(0.04)	(0.04)
Food consumption (log Rs)	0.06*	0.07*	0.03	0.07
	(0.03)	(0.04)	(0.04)	(0.05)
	0.08**	0.09**	0.05	0.08
	(0.03)	(0.04)	(0.04)	(0.05)
	0.08**	0.09*	0.05	0.08
	(0.04)	(0.05)	(0.05)	(0.06)
Meals per day (no.)	0.24**	0.37**	0.25**	0.58**
	(0.10)	(0.16)	(0.13)	(0.27)
	0.29***	0.48***	0.28**	0.66**
	(0.10)	(0.16)	(0.13)	(0.27)
	0.29**	0.48**	0.28**	0.66*
	(0.12)	(0.19)	(0.14)	(0.35)
Outpatient expenses (log Rs)	-0.45**	-0.43*	-0.59**	-0.56*
	(0.23)	(0.23)	(0.26)	(0.31)
	-0.57**	-0.53**	-0.69***	-0.65**
	(0.23)	(0.23)	(0.26)	(0.31)
	-0.57**	-0.53*	-0.69**	-0.65*
	(0.27)	(0.27)	(0.30)	(0.34)
Households financial access (Table 7)				
Savings account (yes/no)	0.19*	0.22*	0.24*	0.26
	(0.10)	(0.11)	(0.13)	(0.18)
	0.27***	0.30***	0.30**	0.32*
	(0.10)	(0.11)	(0.13)	(0.18)
	0.27**	0.30**	0.30**	0.32
	(0.12)	(0.14)	(0.14)	(0.20)
Health insurance (yes/no)	0.17**	0.08	0.25**	0.16**
	(0.07)	(0.05)	(0.12)	(0.07)
	0.20***	0.09*	0.29**	0.18***
	(0.07)	(0.05)	(0.12)	(0.07)
	0.20**	0.09	0.29**	0.18**
	(0.08)	(0.06)	(0.13)	(0.08)
Health care supply (Table 8, 9, and 11)				
Institutional loan (share)	0.010**	0.007*	0.008*	0.007*
	(0.004)	(0.003)	(0.005)	(0.004)
	0.010**	0.008**	0.008	0.007*
	(0.004)	(0.003)	(0.005)	(0.004)
	0.010*	0.008*	0.008	0.007
	(0.005)	(0.004)	(0.006)	(0.005)
Institutional loan (share) - private	0.020**	0.014	0.015*	0.019**
	(0.009)	(0.009)	(0.009)	(0.010)
	0.020**	0.017*	0.014	0.020**
	(0.009)	(0.009)	(0.009)	(0.010)
	0.020*	0.017	0.014	0.020
	(0.012)	(0.011)	(0.011)	(0.012)

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. 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 indicates 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 that is the bias-corrected RD estimator with 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 A5 and A6.

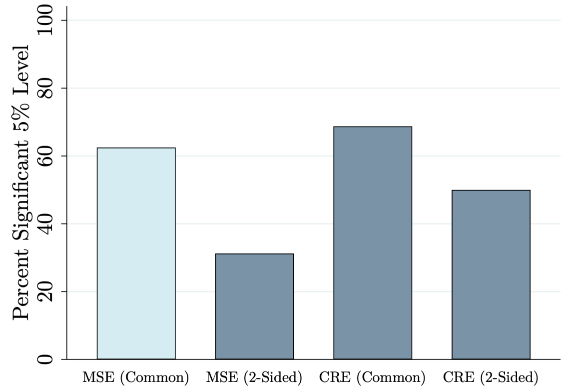
Table A15: **Robustness to Different Bandwidth Selectors: Mechanism Results (2/2)**

	MSE-optimal		CER-optimal	
	Common	Two-sided	Common	Two-sided
	(1)	(2)	(3)	(4)
Health care supply (Table 8, 9, and 11)				
Hospitals (log no.)	0.88***	0.78**	1.32***	1.16**
	(0.33)	(0.33)	(0.46)	(0.49)
	1.14***	1.03***	1.51***	1.34***
	(0.33)	(0.33)	(0.46)	(0.49)
	1.14***	1.03**	1.51***	1.34**
	(0.40)	(0.40)	(0.51)	(0.54)
Hospitals (log no.) - private	0.84**	0.81**	1.28***	1.12**
	(0.36)	(0.36)	(0.49)	(0.52)
	1.11***	1.13***	1.47***	1.33**
	(0.36)	(0.36)	(0.49)	(0.52)
	1.11**	1.13**	1.47***	1.33**
	(0.43)	(0.44)	(0.55)	(0.57)
Hospitals (log no.) - government	0.64*	0.61**	0.91**	0.78*
	(0.33)	(0.30)	(0.45)	(0.42)
	0.78**	0.68**	1.01**	0.84**
	(0.33)	(0.30)	(0.45)	(0.42)
	0.78*	0.68*	1.01**	0.84*
	(0.40)	(0.37)	(0.50)	(0.47)
Survey on problems (Table 10)				
Distance to facility (yes/no)	-0.12***	-0.09**	-0.22***	-0.13***
	(0.04)	(0.04)	(0.07)	(0.05)
	-0.15***	-0.12***	-0.22***	-0.15***
	(0.04)	(0.04)	(0.07)	(0.05)
	-0.15***	-0.12**	-0.22**	-0.15***
	(0.05)	(0.05)	(0.10)	(0.06)
Taking transport to facility (yes/no)	-0.11***	-0.08**	-0.20***	-0.10**
	(0.04)	(0.04)	(0.06)	(0.05)
	-0.15***	-0.12***	-0.20***	-0.13**
	(0.04)	(0.04)	(0.06)	(0.05)
	-0.15***	-0.12***	-0.20**	-0.13**
	(0.05)	(0.05)	(0.09)	(0.06)
No personnel at facility (yes/no)	-0.14**	-0.14**	-0.22***	-0.17**
	(0.06)	(0.06)	(0.08)	(0.07)
	-0.17***	-0.18***	-0.22***	-0.19***
	(0.06)	(0.06)	(0.08)	(0.07)
	-0.17**	-0.18**	-0.22**	-0.19**
	(0.07)	(0.07)	(0.10)	(0.08)
No female personnel at facility (yes/no)	-0.20**	-0.17***	-0.31***	-0.24***
	(0.08)	(0.06)	(0.11)	(0.07)
	-0.24***	-0.21***	-0.32***	-0.27***
	(0.08)	(0.06)	(0.11)	(0.07)
	-0.24***	-0.21***	-0.32**	-0.27***
	(0.09)	(0.07)	(0.14)	(0.09)
No drugs at facility (yes/no)	-0.15**	-0.17***	-0.20**	-0.15**
	(0.07)	(0.06)	(0.08)	(0.07)
	-0.18***	-0.23***	-0.19**	-0.18**
	(0.07)	(0.06)	(0.08)	(0.07)
	-0.18**	-0.23***	-0.19*	-0.18**
	(0.08)	(0.08)	(0.11)	(0.08)

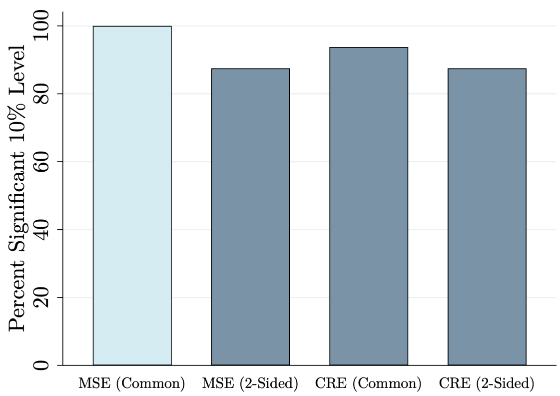
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. The first and second columns are MSE-optimal bandwidths, initially identical and then different to the left and right of the cutoff. The third and fourth columns indicates CER (Coverage Error Rate)-optimal bandwidths, first identical 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 that is the bias-corrected RD estimator with 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 A5 and A6.



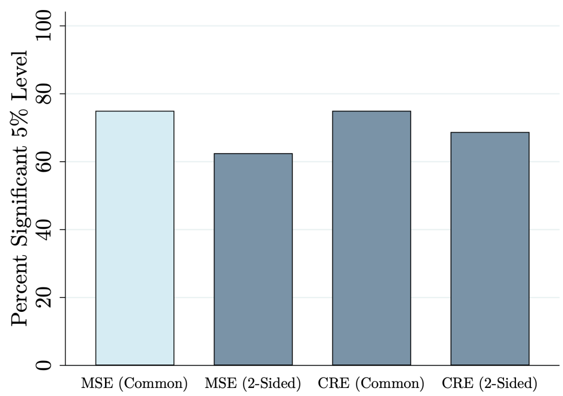
(a) Main Results



(b) Main Results

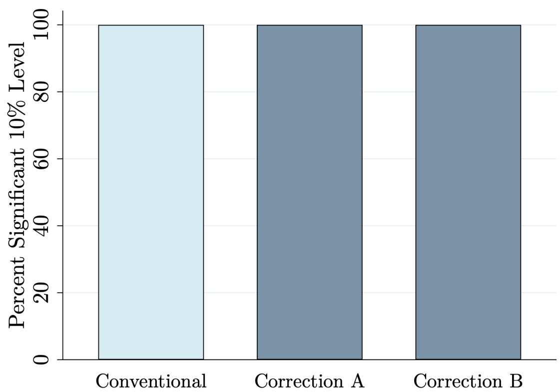


(c) Mechanisms Results

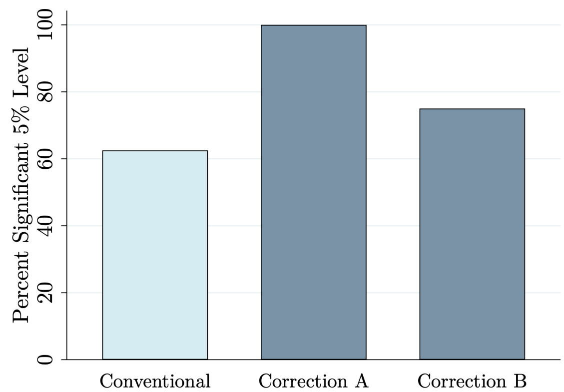


(d) Mechanisms Results

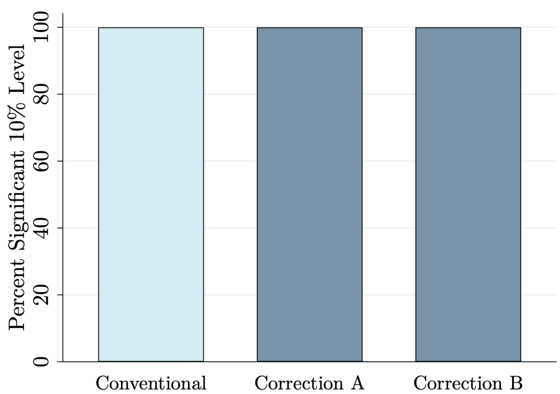
Figure A5. Percent of Results Significant Under Different Bandwidth Selectors. Light blue indicates the main bandwidth (MSE-optimal with common bandwidth to the left and to the right of the cutoff), dark blue indicates alternative bandwidths. The second columns indicate MSE-optimal bandwidths different 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 Table A12, A13, A14, and A15.



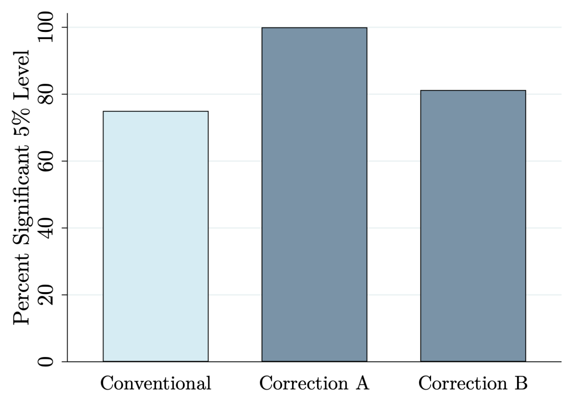
(a) Main Results



(b) Main Results



(c) Mechanisms Results



(d) Mechanisms Results

Figure A6. Percent of Results Significant Under Conventional Specifications and Corrections. Light blue indicates the conventional RD estimator with conventional variance estimator. Correction A is the bias-corrected RD estimator with conventional variance estimator. Correction B is the bias-corrected RD estimator with robust variance estimator (Calonico et al., 2014). Refers to Table A12, A13, A14, and A15.

Table A16: Robustness to Different Polynomial Degrees: Main Results

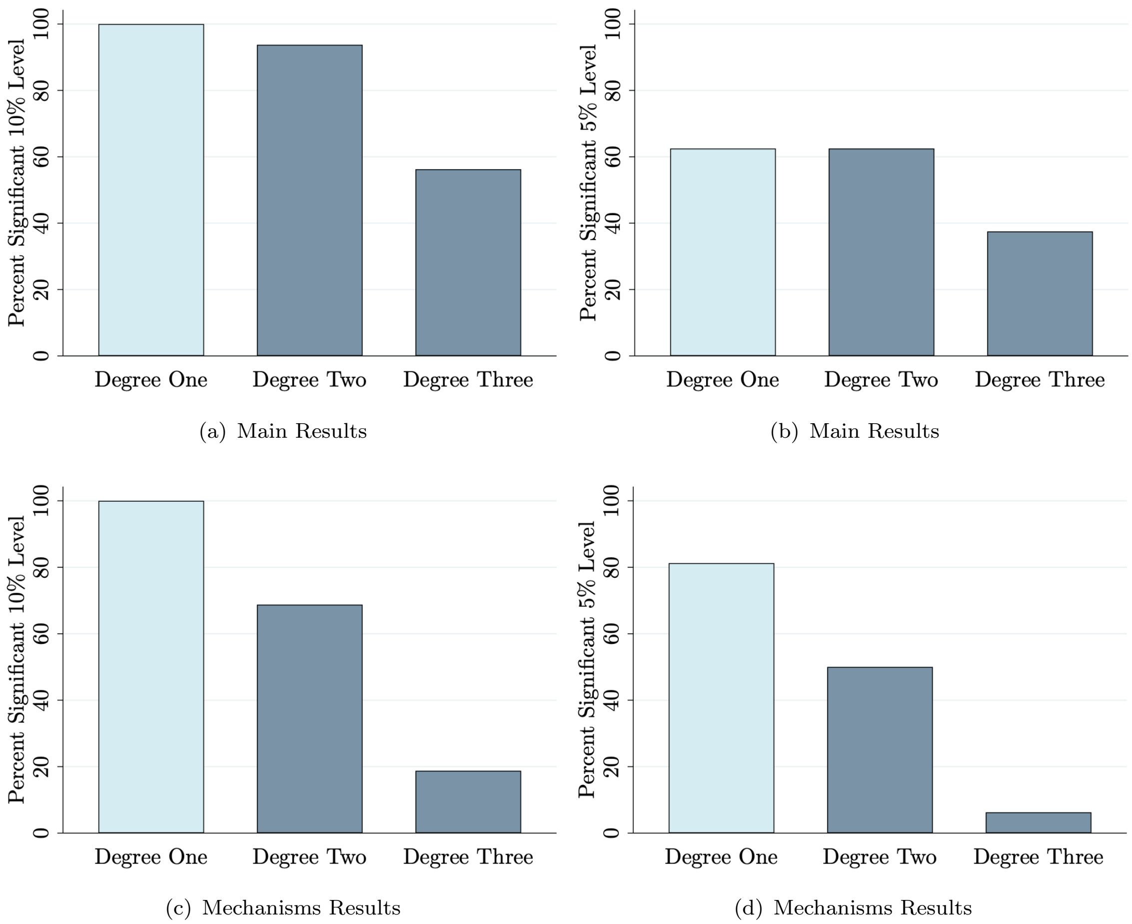
	Polynomial Degree		
	One (1)	**Two** (2)	**Three** (3)
Banks (Table 2)			
Branch licenses 2010 (log no.)	0.19***	0.33***	0.46***
	(0.05)	(0.09)	(0.14)
Branches 2010 (log no.)	0.17***	0.31***	0.44***
	(0.06)	(0.09)	(0.14)
Households health (Table 3)			
Days ill (yes/no)	-0.19**	-0.22*	-0.23
	(0.09)	(0.13)	(0.17)
Days ill (log no.)	-0.29**	-0.35*	-0.41
	(0.12)	(0.19)	(0.26)
Days missed (yes/no)	-0.30***	-0.39**	-0.45*
	(0.10)	(0.17)	(0.24)
Days missed (log no.)	-0.44***	-0.56**	-0.65**
	(0.13)	(0.22)	(0.31)
Medical expenses (yes/no)	-0.18**	-0.21*	-0.23
	(0.08)	(0.12)	(0.16)
Medical expenses (log Rs)	-0.88**	-1.02**	-1.28
	(0.35)	(0.50)	(0.83)
Households health (Table 4)			
Vaccinated child (yes/no)	0.07*	0.16**	0.21**
	(0.04)	(0.08)	(0.10)
Sick child (yes/no)	-0.06*	-0.08	-0.23**
	(0.03)	(0.05)	(0.11)
Health care visit (any reason) (yes/no)	-0.08**	-0.13**	-0.22*
	(0.03)	(0.05)	(0.11)
Health care visit (child's treatment) (yes/no)	-0.02*	-0.06**	-0.13**
	(0.01)	(0.03)	(0.05)
Households health (Table 5)			
Health care facility delivery (yes/no)	0.005***	0.006**	0.003
	(0.002)	(0.003)	(0.005)
Experienced miscarriage (yes/no)	-0.010*	-0.017*	-0.029
	(0.006)	(0.010)	(0.018)
Experienced stillbirth (yes/no)	-0.002*	-0.003*	-0.006
	(0.001)	(0.002)	(0.004)
HC visit (woman's treatment) (yes/no)	-0.051*	-0.106**	-0.175*
	(0.027)	(0.046)	(0.098)

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. For details of the regression refer to the respective main table. Summarized in Figure A7.

Table A17: **Robustness to Different Polynomial Degrees: Mechanism Results**

	Polynomial Degree		
	One	Two	Three
	(1)	(2)	(3)
Households consumption (Table 6)			
Total consumption (log Rs)	0.07**	0.14**	0.19***
	(0.04)	(0.05)	(0.07)
Food consumption (log Rs)	0.06*	0.11***	0.14***
	(0.03)	(0.04)	(0.05)
Meals per day (no.)	0.24**	0.30	0.38
	(0.10)	(0.19)	(0.24)
Outpatient expenses (log Rs)	-0.45*	-0.71	-0.63
	(0.23)	(0.60)	(0.73)
Households financial access (Table 7)			
Savings account (yes/no)	0.22**	0.34*	0.36
	(0.09)	(0.17)	(0.22)
Health insurance (yes/no)	0.15**	0.13	0.20
	(0.07)	(0.08)	(0.14)
Health care supply (Table 8, 9, and 11)			
Institutional loan (share)	0.010**	0.012*	0.015
	(0.004)	(0.006)	(0.011)
Institutional loan (share) - private	0.020**	0.027*	0.033
	(0.009)	(0.014)	(0.023)
Hospitals (log no.)	0.88***	1.22**	1.74
	(0.33)	(0.56)	(1.25)
Hospitals (log no.) - private	0.84**	1.44**	1.92
	(0.36)	(0.69)	(1.54)
Hospitals (log no.) - government	0.64*	0.99	1.49
	(0.33)	(0.62)	(1.35)
Survey on problems (Table 10)			
Distance to facility (yes/no)	-0.12***	-0.10	-0.28
	(0.04)	(0.07)	(0.17)
Taking transport to facility (yes/no)	-0.11***	-0.11**	-0.24
	(0.04)	(0.05)	(0.16)
No personnel at facility (yes/no)	-0.14**	-0.23**	-0.28
	(0.06)	(0.10)	(0.22)
No female personnel at facility (yes/no)	-0.20**	-0.30**	-0.45*
	(0.08)	(0.12)	(0.23)
No drugs at facility (yes/no)	-0.15**	-0.22**	-0.31
	(0.07)	(0.11)	(0.23)

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. For details of the regression refer to the respective main table. Summarized in Figure A7.



(a) Main Results

(b) Main Results

(c) Mechanisms Results

(d) Mechanisms Results

Figure A7. Percent of Results 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 A16 and A17.

Table A18: Placebo Cutoffs: Main Results

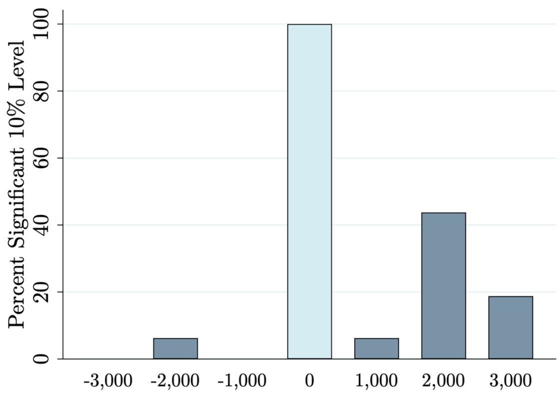
	Placebo cutoff						
	-3,000	**-2,000**	**-1,000**	**0**	**1,000**	**2,000**	**3,000**
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Banks (Table 2)							
Branch licenses (log no.)	0.92	0.01	0.22	0.00	0.78	0.06	0.04
Branches (log no.)	0.87	0.52	0.40	0.00	0.50	0.14	0.04
Households health (Table 3)							
Days ill (yes/no)	0.70	0.98	0.33	0.03	0.26	0.03	0.55
Days ill (log no.)	0.84	0.95	0.33	0.02	0.18	0.04	0.62
Days missed (yes/no)	0.99	0.19	0.58	0.00	0.33	0.06	0.51
Days missed (log no.)	0.93	0.44	0.46	0.00	0.32	0.06	0.58
Medical expenses (yes/no)	0.90	0.91	0.80	0.02	0.24	0.02	0.62
Medical expenses (log Rs)	0.68	0.53	0.93	0.01	0.27	0.11	0.28
Households health (Table 4)							
Vaccinated child (yes/no)	0.50	0.55	0.65	0.07	0.27	0.07	0.49
Sick child (yes/no)	0.21	0.64	0.96	0.06	0.15	0.70	0.37
HC visit (any reason) (yes/no)	0.57	0.78	0.79	0.02	0.29	0.87	0.56
HC visit (child's treatment)	0.44	0.84	0.56	0.10	0.35	0.99	0.44
Households health (Table 5)							
Health care facility delivery (yes/no)	0.37	0.38	0.99	0.01	0.86	0.84	0.51
Experienced miscarriage (yes/no)	0.59	0.19	0.67	0.09	0.27	0.47	0.45
Experienced stillbirth (yes/no)	0.38	0.92	0.83	0.09	0.00	0.34	0.09
HC visit (woman's treatment) (yes/no)	0.46	0.34	0.84	0.06	0.51	0.65	0.87

P-values of respective regressions with different (placebo) cutoffs shown. For details of the regressions refer to the respective main table. Summarized in Figure A8.

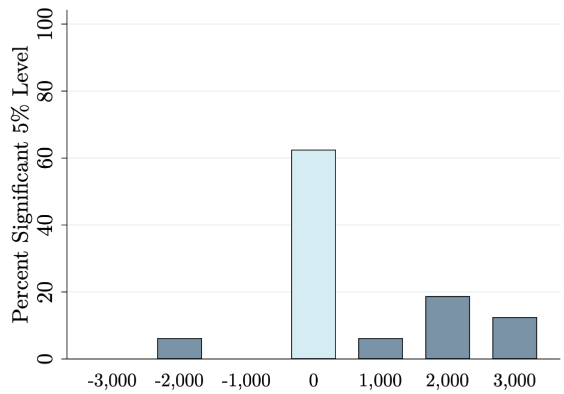
Table A19: Placebo Cutoffs: Mechanism Results

	Placebo cutoff						
	-3,000	**-2,000**	**-1,000**	**0**	**1,000**	**2,000**	**3,000**
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Households consumption (Table 6)							
Total consumption (log Rs)	0.51	0.91	0.16	0.04	0.93	0.64	0.25
Food consumption (log Rs)	0.82	0.52	0.38	0.10	0.27	0.67	0.76
Meals per day (no.)	0.73	0.53	0.26	0.02	0.98	0.91	0.48
Outpatient expenses (log Rs)	0.59	0.25	0.27	0.05	0.71	0.17	0.16
Households financial access (Table 7)							
Savings account (yes/no)	0.57	0.84	0.24	0.01	0.46	0.24	0.77
Health insurance (yes/no)	0.56	0.06	0.53	0.03	0.91	0.04	0.74
Health care supply (Table 8, 9, and 11)							
Institutional loan (share)	0.77	0.90	0.74	0.02	0.23	0.90	-
Institutional loan (share) - private	0.69	0.65	0.25	0.03	0.05	0.93	0.85
Hospitals (log no.)	0.59	0.97	0.01	0.01	0.17	0.91	0.87
Hospitals (log no.) - private	0.50	0.85	0.03	0.02	0.30	0.53	0.62
Hospitals (log no.) - government	0.72	0.14	0.36	0.05	0.41	0.21	0.19
Survey on problems (Table 10)							
Distance to facility (yes/no)	0.86	0.66	0.71	0.01	0.29	0.70	0.00
Taking transport to facility (yes/no)	0.86	0.38	0.61	0.00	0.79	0.35	0.09
No personnel at facility (yes/no)	-	0.33	0.34	0.02	0.57	0.92	-
No female personnel at facility (yes/no)	0.82	0.77	0.25	0.01	0.42	0.95	-
No drugs at facility (yes/no)	0.75	0.38	0.50	0.03	0.42	0.45	-

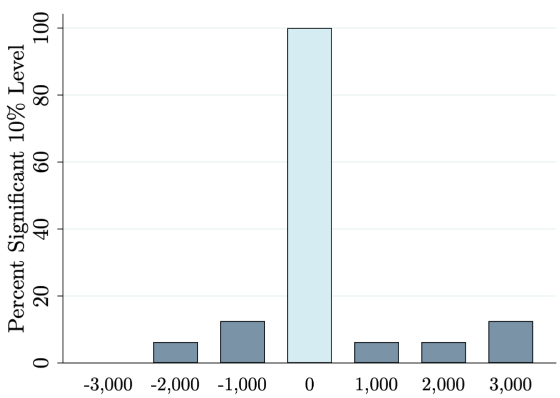
P-values of respective regressions with different (placebo) cutoffs shown. For details of the regressions refer to the respective main table. Summarized in Figure A8.



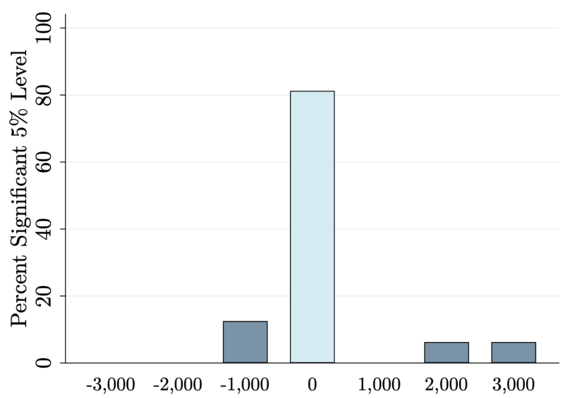
(a) Main Results



(b) Main Results



(c) Mechanisms Results



(d) Mechanisms Results

Figure A8. Percent of Results 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 A18 and A19.

Table A20: Standard Error Adjustments: Main Results

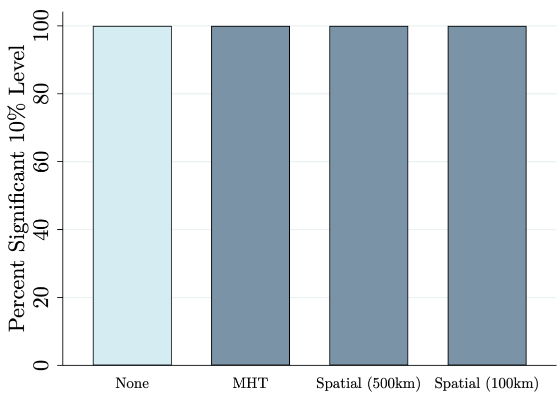
	Adjustment			
	None	**Multiple hypothesis testing**	**Spatial correlation (500km)**	**Spatial correlation (100km)**
	(1)	(2)	(3)	(4)
Banks (Table 2)				
Branch licenses 2010 (log no.)	0.00	0.01	0.00	0.00
Branches (log no.)	0.00	0.02	0.00	0.00
Households health (Table 3)				
Days ill (yes/no)	0.03	0.04	-	-
Days ill (log no.)	0.02	0.03	-	-
Days missed (yes/no)	0.00	0.02	-	-
Days missed (log no.)	0.00	0.02	-	-
Medical expenses (yes/no)	0.02	0.03	-	-
Medical expenses (log Rs)	0.01	0.03	-	-
Households health (Table 4)				
Vaccinated child (yes/no)	0.07	0.04	-	-
Sick child (yes/no)	0.06	0.04	-	-
HC visit (any reason) (yes/no)	0.02	0.03	-	-
HC visit (child's treatment)	0.10	0.05	-	-
Households health (Table 5)				
Health care facility delivery (yes/no)	0.01	0.03	-	-
Experienced miscarriage (yes/no)	0.09	0.05	-	-
Experienced stillbirth (yes/no)	0.09	0.05	-	-
HC visit (woman's treatment) (yes/no)	0.06	0.04	-	-

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. For details of the regression refer to the respective main table.

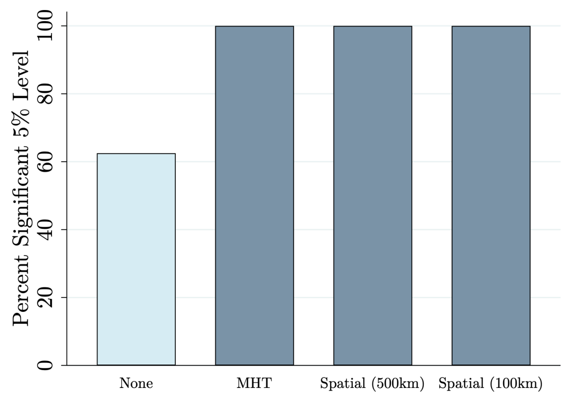
Table A21: Standard Error Adjustments: Mechanism Results

	Adjustment			
	None	**Multiple hypothesis testing**	**Spatial correlation (500km)**	**Spatial correlation (100km)**
	(1)	(2)	(3)	(4)
Households consumption (Table 6)				
Total consumption (log Rs)	0.04	0.04	-	-
Food consumptino (log Rs)	0.10	0.05	-	-
Meals per day (no.)	0.02	0.03	-	-
Outpatient expenses (log Rs)	0.05	0.04	-	-
Households financial access (Table 7)				
Savings account (yes/no)	0.05	0.04	-	-
Health insurance (yes/no)	0.02	0.03	-	-
Health care supply (Table 8, 9, and 11)				
Institutional loan (share)	0.02	0.03	0.08	0.01
Institutional loan (share) - private	0.03	0.04	0.09	0.02
Hospitals (log no.)	0.01	0.03	0.04	0.03
Hospitals (log no.) - private	0.02	0.03	0.10	0.04
Hospitals (log no.) - government	0.05	0.04	0.00	0.04
Survey on problems (Table 10)				
Distance to facility (yes/no)	0.01	0.03	-	-
Taking transport to facility (yes/no)	0.00	0.02	-	-
No personnel at facility (yes/no)	0.02	0.03	-	-
No female personnel at facility (yes/no)	0.01	0.03	-	-
No drugs at facility (yes/no)	0.03	0.04	-	-

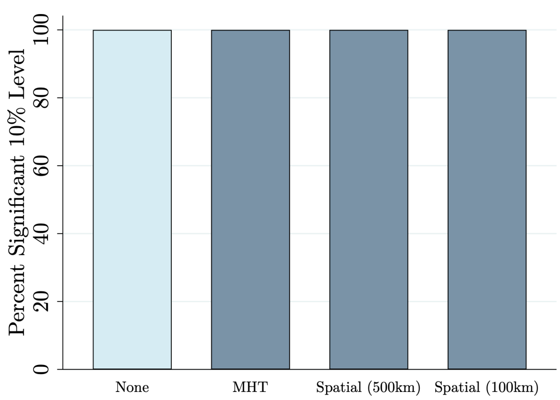
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. For details of the regression refer to the respective main table.



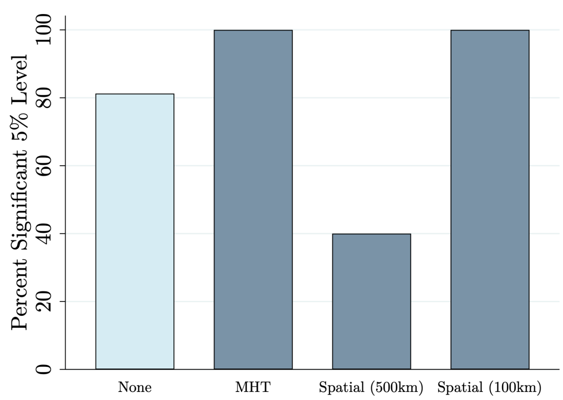
(a) Main Results



(b) Main Results

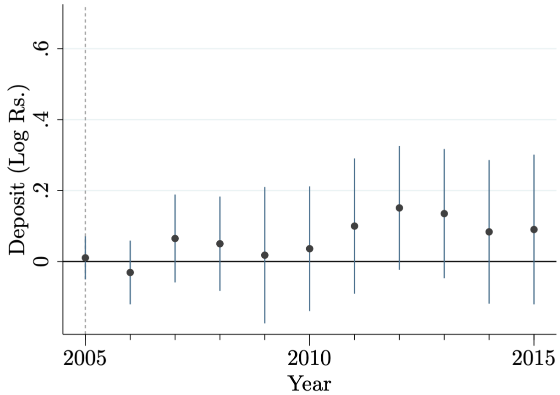


(c) Mechanisms Results

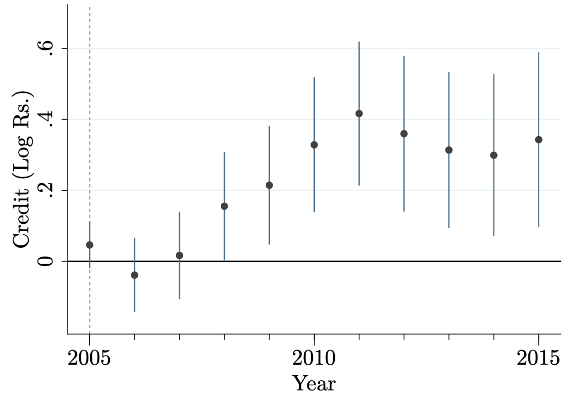


(d) Mechanisms Results

Figure A9. Percent of Results 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 (family-wise error rate), Column 3 and 4 to spatial correlation (Conley standard errors, 100km and 500km). Refers to Table A20 and A21.



(a) Bank Deposit (Dynamics)



(b) Bank Credit (Dynamics)

Figure A10. Dynamics of Aggregated Deposit and Credit for Private Banks. Data RBI. The dynamic figures depict coefficients on aggregated deposit and credit amounts reported. Credit and deposit are measured in billion rupees. The focus is on private banks, which experienced a particularly large growth post 2005. Coefficients for all banks are insignificant.

Discussion A3. One question that arises is whether it is profitable for banks to open branches in underbanked districts. Answering this question requires data on branch profits on at least the district level. 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. However, it is possible to make one specific statement: 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 to be profitable for banks.

Table A22: Smoothness Pre-Policy: Financial Access

	Any loan (yes/no) (1)	Largest loan amount (log Rs) (2)	Largest loan from bank (yes/no) (3)	Health insurance (yes/no) (4)
Treated	0.00	0.12	-0.00	0.01
	(0.10)	(0.86)	(0.03)	(0.01)
Control Mean	0.45	3.03	0.11	0.02
Change (%)	0.24	13.22	-2.81	55.55
First Stage	0.69	0.69	0.71	0.68
Bandwidth	2,950	2,947	4,322	3,086
Efficient Obs.	16,402	14,893	21,224	16,057
Observations	36,913	33,825	37,052	35,204
Baseline Control	No	No	No	No

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. Data IHDS I (2004/2005). Variable in Rs is transformed to log and trimmed at the 10th and 90th percentile.

Table A23: Smoothness Pre-Policy: Morbidity Rate

	Morbidity		Economic consequences			
	Days ill		Days missed		Medical expenses	
	(yes/no) (1)	(log no.) (2)	(yes/no) (3)	(log no.) (4)	(yes/no) (5)	(log Rs) (6)
Treated	-0.06	-0.11	-0.11	-0.19	-0.08	-0.14
	(0.06)	(0.13)	(0.08)	(0.14)	(0.06)	(0.27)
Control Mean	0.42	0.64	0.34	0.48	0.41	1.32
Change (%)	-14.77	-10.49	-31.90	-17.68	-18.73	-13.03
First Stage	0.71	0.69	0.67	0.67	0.72	0.69
Bandwidth	4,432	3,418	2,797	2,524	4,580	3,566
Efficient Obs.	20,799	15,574	14,730	12,122	21,585	16,019
Observations	35,480	31,375	35,294	32,442	35,480	31,812
Baseline Control	No	No	No	No	No	No

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. Data IHDS I (2004/2005). Household level. All variables measured in currency Rs are in log form and trimmed at the 10% and 90% level. All illnesses refer to fever, diarrhea, or cough. Days missed measures the number of days that the household was not able to do usual activities and had to miss work or school. All questions refer to the past 30 days.

Table A24: Baseline Control: Morbidity Rate Decreases

	Morbidity		Economic consequences			
	Days ill		Days missed		Medical expenses	
	(yes/no) (1)	(log no.) (2)	(yes/no) (3)	(log no.) (4)	(yes/no) (5)	(log Rs) (6)
Treated	-0.18**	-0.25**	-0.25***	-0.39***	-0.15**	-0.61**
	(0.07)	(0.12)	(0.08)	(0.12)	(0.06)	(0.31)
Control Mean	0.53	0.81	0.40	0.57	0.51	2.05
Change (%)	-34.11	-22.37	-63.27	-32.40	-30.30	-45.60
First Stage	0.67	0.70	0.68	0.70	0.68	0.70
Bandwidth	2,600	2,894	3,237	3,087	3,139	3,867
Efficient Obs.	13,870	14,559	17,374	15,469	17,113	17,435
Observations	35,294	27,603	35,371	29,387	35,294	28,610
Baseline Control	Yes	Yes	Yes	Yes	Yes	Yes

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. Data IHDS II (2011/2012). Household level. All variables measured in currency Rs are in log form and trimmed at the 10% and 90% level. All illnesses refer to fever, diarrhea, or cough. Days missed measures the number of days that the household was not able to do usual activities and had to miss work or school. All questions refer to the past 30 days.

Table A25: Placebo Test: Long-Term Illnesses

	Morbidity	Economic consequences			
	Days ill	Days missed		Medical expenses	
	(yes/no) (1)	(yes/no) (2)	(log no.) (3)	(yes/no) (4)	(log Rs) (5)
Treated	-0.00	-0.05	-0.02	0.00	-0.20
	(0.05)	(0.05)	(0.15)	(0.05)	(0.37)
Control Mean	0.39	0.30	0.59	0.37	1.67
Change (%)	-0.96	-15.55	-1.57	0.02	-17.98
First Stage	0.65	0.62	0.60	0.64	0.58
Bandwidth	2,189	2,038	1,934	2,107	1,920
Efficient Obs.	11,716	9,962	8,697	10,981	8,700
Observations	35,103	34,883	31,426	35,103	31,621
Baseline Control	No	No	No	No	No

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. Data IHDS II (2011/2012). Household level. All variables measured in currency Rs are in log form and trimmed at the 10% and 90% level. All illnesses refer to a variety of long-term diseases including cancer, diabetes, or heart disease. Days missed measures the number of days that the household was not able to do usual activities and had to miss work or school. All questions refer to the past 365 days.

Table A26: 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	-0.10*	0.20***	-0.00
	(0.06)	(0.06)	(0.00)
Control Mean	0.46	0.51	0.00
Change (%)	-21.09	40.20	-7.08
First Stage	0.61	0.60	0.61
Bandwidth	2,020	1,955	2,015
Efficient Obs.	43,390	42,833	43,230
Observations	159,078	159,078	159,078
Baseline Control	No	No	No

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. Data DHS (2015/2016). Household level.

Table A27: Larger Effects with High Probability to Take Up Financial Instruments

	Savings account		Bank loan		Health insurance	
	High	Low	High	Low	High	Low
	Days ill (yes/no)	Days ill (yes/no)	Days ill (yes/no)	Days ill (yes/no)	Days ill (yes/no)	Days ill (yes/no)
	(1)	(2)	(3)	(4)	(5)	(6)
Treated	-0.29**	-0.10*	-0.24**	-0.12**	-0.33***	-0.07
	(0.12)	(0.06)	(0.11)	(0.06)	(0.12)	(0.08)
Control Mean	0.53	0.53	0.53	0.53	0.53	0.56
Change (%)	-55.10	-19.27	-45.52	-23.61	-62.55	-13.31
First Stage	0.57	0.75	0.59	0.73	0.55	0.82
Bandwidth	2,222	2,953	2,226	2,916	2,336	1,718
Efficient Obs.	7,656	5,976	7,608	5,934	7,838	3,506
Observations	23,061	13,739	23,249	13,555	22,687	13,731
Baseline Control	No	No	No	No	No	No

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. Data IHDS II (2011/2012). Household level. I run three predictions, one for each of taking up a savings account (Column 1 and 2), having a bank loan (Column 3 and 4), and having health insurance (Column 5 and 6). All predictions are run with state fixed effects within a bandwidth of -3,000 to +3,000. Baseline characteristics from the IHDS I are whether their largest loan was from a bank, whether they are urban, their assets, and their per capita consumption. The odd columns use the sample of households in the upper half of the respective distribution. The even columns use the sample of households in the lower half of the respective distribution. The outcome is the number of days ill in the IHDS I.