Financial Inclusion and Female Education[*]

Manya Malik[†]

Click here for the latest version

Abstract

I examine the effect of increased access to financial services on girls' education through the lens of a 2005 bank expansion policy in India. The policy provided exogenous variation in bank availability by incentivizing private banks to make inroads into the under-served areas. Labeled as "underbanked", these areas had a population to bank branch ratio greater than the national average. Using a regression discontinuity framework, I find that ten years after the policy, girls in underbanked districts are 3 to 8 percent more likely to be enrolled in school in comparison to girls in non-underbanked districts. The effects vary across age groups, with stronger impacts observed for adolescent girls. I identify higher savings and an income effect as potential channels driving increased enrollment for girls. Additionally, improved labor market outcomes for lower-caste groups suggest that economically and historically disadvantaged households could be driving these results.

Keywords: Banking, Education, Gender

[*]I would like to express my gratitude to Alan Griffith, Rachel Heath, David Bloom and Joseph Dieleman for their invaluable guidance and support throughout the project.I am also thankful to the speakers at Joint Seminar for Development Economics and the participants of Labor and Development brownbag for their helpful comments and insights.

[†]Department of Economics, University of Washington, Seattle. Email: mmalik16@uw.edu

1 Introduction

Financial inclusion plays a pivotal role in economic growth (Rajan & Zingales, 1996) and poverty reduction, particularly in developing nations (Burgess & Pande, 2005; Barboni et al., 2021). Besides the macro-level effects, examining the micro-level impact offers critical insights into how financial services affect different segments of the society, particularly across gender (Ashraf et al.,2010) and income distribution (Banerjee et al.,2019). This paper aims to investigate how access to financial services can serve as a step towards narrowing the gender gap in human capital formation, a key driver of inter-generational mobility (Machin, 2007). By improving access to credit (Kaboski & Townsend, 2011), savings (Brune et al., 2011; Prina 2013), and economic opportunities (Bruhn & Love, 2014), financial inclusion has the potential to alleviate financial constraints, enabling marginalized groups to invest in education and foster upward mobility across generations.

In this study, I examine the impact of bank expansion on educational outcomes in Indian context. To get exogenous variation in bank availability, I use a 2005 bank expansion policy introduced by Reserve Bank of India that incentivized private banks to expand to "underbanked" areas. These areas referred to districts that had a population to bank branch ratio greater than the national average. Since the targeted areas for bank expansion were defined based on an arbitrary cutoff, I can exploit the same in a regression discontinuity setup to determine the causal effect of increased financial access as a result of this policy.

To begin with, I investigate how the banks responded to this policy. Did the incentives lead to a higher presence of brick-and-mortar bank branches in underbanked districts of India? Using RBI Bank Branch Statistics data, I find that private banks respond with higher growth in these areas whereas there is no effect in public sector banks. Within seven years of the policy, underbanked districts have 27 percent more private bank openings than control districts. In terms of the total number of private banks, underbanked districts have around 36 percent more bank branches that control districts in 2010.

Having established a direct effect of this policy on the targeted banking outcomes, the next step is to see if there are downstream effects on educational outcomes in the long run. To examine this ten years after the policy was introduced, I use nationally representative household level ASER survey data that records schooling outcomes for 5-16 year olds. Results show that girls in underbanked districts are 3.4 percentage points more likely to be enrolled in school. Upon a heterogeneity analysis by age, I find this effect to be higher for older girls. In the age group of 15-16 years, girls in underbanked districts are 6.6 percentage points more likely to be in school in comparison to girls in non-underbanked districts. This corresponds to a 7.6 percent increase over the average enrollment for girls in non-underbanked districts. In contrast, I find insignificant and small effects for boys.

To interpret these results as causal effect of bank expansion under a regression discontinuity setup, I provide evidence that there is no manipulation in treatment status around the cutoff. Since the underbanked assignment was based on population to bank ratio constructed using 2001 population data, it is unlikely that the districts anticipated this policy in 2001 and manipulated district population in response. The denominator includes sum of all types of banks operated by different enterprises which would be very hard to manipulate. I also run the McCrary(2008) density test to check for smoothness in number of districts around the cutoff. Additionally, I check for pre-policy balance to provide evidence that treatment and control districts around the cutoff are similar. Lastly, I perform robustness checks by using placebo cutoffs, excluding observations very close to the cutoff and using a polynomial of degree two.

In addition to the primary enrollment outcomes, I also investigate several channels through which banking access could be affecting education. The first question to ask would be, how are households using this increased access to banking? Findings suggest that the policy doesn't lead to an increase in credit uptake but the likelihood of saving in a post office account or savings scheme increases by almost 85 percent. So, there is evidence on higher use of financial products in the form of savings and not higher borrowing. By studying the long term impact of this policy, I can also examine how households benefit from general equilibrium effects. Higher business activity in the underbanked districts leading to an increase in labor force participation and wages affects household income positively. When I distinguish households by caste and examine labor market outcomes, I find that these effects are driven by historically disadvantaged groups of SC, ST and OBC households. In India, these groups are usually poorer and have stronger cultural norms resulting in a higher gender gap in secondary education. This suggests that the positive effects of bank expansion on girls' education are driven by socio-economically disadvantaged households who are most likely to be on the margin. Through an income effect and savings, girls in these marginalized households are likely to stay longer in school.

This paper aligns with the existing literature on formal banking access in India. Previous studies have documented the effects of bank expansion during the social banking era in 1970s-1980s and its impact on poverty reduction (Burgess and Pande , 2005; Kochar, 2011). In context to the RBI bank expansion policy used in this paper , Young 2017 finds evidence of positive effects on agricultural productivity and local GDP growth using nightlights data. Other studies look at firm level dynamics (Ritadhi and Mahajan, 2022) and focus on growth in manufacturing firms (Jao and Mo, 2023). The policy also had implications on labor market outcomes leading to labor reallocation towards the manufacturing sector and higher wage earnings (Jakaria 2023). Cramer (2021) is the only study that looks at the impact on household well-being and finds positive effects on health through an increase in healthcare supply. My paper aligns with this literature focusing on household well-being but differs along three dimensions. First, I study education and how female education can be more responsive to financial deepening. This is the first paper that draws attention to this aspect of household well-being using the RBI policy. Second, I explore heterogeneity in effects by studying how the effects vary across age groups and look at households distinguished by caste. My analysis sheds light on whether the positive effects were driven by the groups that are economically and historically disadvantaged in India. Third, by examining educational outcomes 10 years after the policy was introduced, I am able to discern the long term effects of this policy on households, which takes into account the general equilibrium effects of bank expansion as well. While I am not able to disentangle the effects of increased savings and increased income through labor market, it is evident that financial access can affect households through multiple channels in the long run.

This paper also contributes to the empirical literature connecting financial access and household well-being , either through quasi-experimental approaches (Agarwal et al, 2017; Bruhn and Love ,2014; Patnama and Yao, 2020; Fonseca and Matray ,2022) or through randomized control trials (Attanasio et al.,2015; Schaner 2016;Banerjee et al.,2015). While most RCTs find small or no effects (Banerjee et al.,2015; Tarozzi et al.,2015; Dupas et al., 2018) on education, they can primarily look at short term effects of credit expansion or savings. My paper differs from these studies by considering a longer time frame that allows for general equilibrium effects to kick in.

The results in my analysis have considerable policy implications. Financial deepening can not only have direct effects on access to credit or financial products, but it could also lead to structural changes in the labor market through an effect on firms and households. The findings suggest that financial system can influence a household's decision to invest in education, particularly for girls. A policy aimed at increasing access to finance can potentially help address the major challenge of closing gender gap in secondary education, especially within groups where financial constraints and cultural norms are stronger.

The remainder of this paper is organized as follows: Section 2 introduces the Reserve Bank of India's (RBI) bank expansion policy. Section 3 describes the data, while Section 4 outlines the identification strategy. Section 5 presents the main results, focusing on the effects on banking and education. Section 6 provides robustness checks, and Section 7 explores potential channels. Finally, Section 8 concludes.

2 Policy Context

The policy I use in this paper for exogenous variation in bank availability was introduced by Reserve Bank of India in September 2005. According to the Branch Authorization Policy (BAP), private banks were incentivized to expand operations by opening more branches in "underbanked" areas. These areas refer to the districts which had fewer existing branches or low competition.

To formally define a district as "underbanked", RBI leveraged the density ratio i.e. population to bank branch ratio for a district. Districts for which this density ratio exceeded the national average were classified as "underbanked" by RBI.

$$\text{Underbanked}_{District} = 1 \text{ if } \frac{\text{Population}_{District}}{\text{Number of Bank Branches}_{District}} > \frac{\text{Population}_{National}}{\text{Number of Bank Branches}_{National}}$$

In order to construct this measure, RBI used population data from the 2001 census and RBI data on the total number of bank branches in operation in 2005. A list of underbanked districts was released by RBI in order to facilitate banks to identify them. Since this list doesn't include the corresponding density ratio or the national average cutoff used by RBI, I reconstruct these numbers using raw Census and RBI data.

This policy also simplified the underlying process of obtaining license to open new branches through two primary procedural changes. First, the existing system of authorizing individual branches on a case by case basis was replaced by a system of approving on an annual basis. The banks were granted flexibility to submit their annual branch expansion plan to RBI at any time during the year. Second, the process was made more centralized by allowing banks to approach RBI directly, instead of having to approach regional offices under the old system. To incentivize banks towards underbanked districts, the policy stated that banks could increase their chances of obtaining licenses for their preferred locations if they opened more branches in districts listed as underbanked by RBI. Hence, by penetrating into these districts, banks could increase their chances of getting favorable treatment from RBI with regard to their overall expansion plans.

While this policy indicates "incentivization", it didn't state any explicit rules unlike some of the other policies that have existed in the Indian financial landscape in the past. From 1969-1990, bank expansion was mainly led by public sector banks, focussing heavily on areas without any banks. In that case, expansion policy was ratio based where a bank had to open a specified number of offices [1] in unbanked rural centres for every new office in urban centres. However, following liberalization in 1991, banks were given freedom to expand based on business potential and financial viability, which continued until 2005.

In this paper, I look at increasing financial inclusion through the lens of BAP policy, with the primary focus on increasing the presence of brick-and-mortar bank branches in underbanked districts. In the last decade, Indian financial system has become much more sophisticated, such that the delivery of banking services is through a combination of physical branches as well as branchless models like ATMs or mobile banking. Future research could look at how a mix of these models evolved in underbanked areas and affected the economy.

3 Data

This section describes multiple data sources used in my analysis. Since RBI published a list of underbanked districts, without actually releasing data on the underlying density ratio and national average cutoff used, I start by reconstructing these variables. I use Census 2001 data to get population for every district in India. For data on district wise number of bank branches, I refer to Bank Branch Statistics on RBI data portal [2]. The earliest available data is from the first quarter of 2006, which I use to construct the density ratio. I calculate the national average cutoff as 14,976 (population per bank branch) and

[1] Under the norm introduced in 1968, this ratio was 1:1 between banked and unbanked centres. Starting 1970, licenses were based on a ratio of 1:2 between urban centre and rural or semi-urban centres for public banks and 1:3 for other banks.From 1977, this was increased to 1:4 for all banks.

[2] RBI Data Portal

use the computed density ratio as a running variable in my RD setup.

Out of a total of 593 districts existing in 2001, RBI classified 375 as underbanked. Based on the reconstruction, I classify all except seven districts correctly. As a result, I use a fuzzy RD setup instead of sharp to account for the fact that the cutoff doesn't perfectly predict the treatment status.

For the first stage outcomes, I look at two variables from RBI banking data : total number of bank branches and the new bank branches opened annually. I use data on these two variables for private banks and public sector banks between 2005 - 2012. This allows me to look at pre-policy balance in banking as well as dynamic effects across different years.

Next, I look at IHDS dataset from 2011-2012 for household level banking variables. This includes information on borrowing from formal and informal sources, savings and investment in different assets and information on the purpose of loan. District level data on bank branches in combination with household level bank utilization data will allow me to observe how effective this policy was in pushing financial inclusion.

For the second stage outcomes related to education, I use two nationally representative household level surveys. First is the Annual Status of Education Report (ASER) from 2015-2016. This survey provides information on school enrollment as well as certain household characteristics such as mother and father's education. I use this survey for my primary outcomes on school enrollment by gender and age. To establish pre-policy balance, I use ASER 2007 survey which is the earliest possible round available.

I complement ASER household survey with NSSO Education Survey from 2014-2015. This is the NSS 71st round devoted to the subject of social consumption, specifically education. It includes detailed information on expenditure incurred on the education of household members and reasons for dropout. I use this to provide suggestive evidence of increased expenditure on education by households in underbanked districts. Lastly, I also use NSS 68th round Employment and Unemployment Survey to examine labor market outcomes for all households and households distinguished by caste. Summary statistics have been provided in Appendix.

I use multiple datasets to gather data on population, banking, employment and education which are spread out across multiple years. Between 2001 to 2016, the total number of districts in India increased from 593 to 620, along with changes in district boundaries. To account for this, I collapse all my data to 2001 census district level and map the new districts to old ones to the best of my knowledge. This could be one of the reasons why the national average used in this paper differs from the one used by Cramer (2021). Despite using a different national average cutoff, my reconstruction method assigns only seven districts a status that is different from the one assigned by RBI, as opposed to ten districts in Cramer(2021).

4 Empirical Methodology

To identify the causal impact of bank expansion on educational outcomes, I will be relying on the regression discontinuity design. Since RBI uses an arbitrary cutoff of national average of population to bank branch ratio for classification of districts, I can exploit the same in an RD design. Based on this classification, treatment can be defined as "underbanked" i.e. the treatment group should consist of all districts with a density ratio above the national average. Similarly, control group should include all districts with a ratio below national average. Density ratio for each district is used as the running variable in my analysis.

One important consideration is that RBI policy documents specify the exact procedure but do not provide the district wise density ratio or the value of national average used. As a result, I reconstructed this data based on the 2001 population census and the RBI data on the number of bank branches existing in March 2006. Based on the reconstructed numbers and a national average cutoff of 14976, I assign the districts as underbanked or not. For 7 out of 593 districts, the assignment using my methodology differs from the status given by RBI. There are about three districts that I assign as underbanked but are not present in the RBI underbank list whereas four districts that are in the RBI underbank list but should not be, based on my methodology[3]. This could be due to different data sources used for number of bank branches. RBI documents do not specify the exact file that was used for the number of bank branches, and since the data is regularly updated, I refer to the oldest possible version of the document. Secondly, it is plausible that RBI exercised discretion when assigning the underbank status for some of the districts. Hence, the rule based on national average cutoff wouldn't determine treatment perfectly for all districts.

Keeping this in mind, I use a fuzzy RD setup instead of a sharp RD. The national average cutoff determines the districts that are eligible to be assigned as underbanked, but the assignment rule is probabilistic. Given that the density ratio goes from below national average to above national average, the probability of being classified as underbanked does not go from 0 to 1.

The identification equation used is similar to Cramer(2021). For unit i, in district d and state s :

$$\begin{aligned} \text{Underbanked}_{ds} = \alpha + \alpha_1 \text{AboveCutoff}_{ds} + \alpha_2 \text{Ratio}_{ds} + \\ \alpha_3 \text{Ratio}_{ds} * \text{AboveCutoff}_{ds} + \delta_s + \epsilon_{ids} \end{aligned} \tag{1}$$

$$\begin{aligned} \text{Y}_{ids} = \beta + \beta_1 \text{Underbanked}_{ds} + \beta_2 \text{Ratio}_{ds} + \\ \beta_3 \text{Ratio}_{ds} * \text{AboveCutoff}_{ds} + \delta_s + \epsilon_{ids} \end{aligned} \tag{2}$$

Here, *Underbanked* refers to a dummy variable which takes a value of 1 if a district was assigned the underbanked status by RBI. *AboveCutoff* refers to a dummy variable which is equal to 1 if the density ratio of a district is above the national average cutoff. *Ratio* refers to the value of density ratio. Y is the outcome variable that is defined at the individual or household level.

In equation (1), α_1 gives us the first stage coefficient i.e. by how much does the prob-

[3]Kamrup, Puri, Ujjain, Mahesana are in the RBI underbanked list but have a density ratio below the national average cutoff. Pulwama, Nayagarh and Patan have a density ratio above the national average but are not included in the RBI list.

ability of being assigned as underbanked change when the density ratio changes from below to above the cutoff level. Given a fuzzy design, it doesn't jump from 0 to 1 at the cutoff but is relatively strong around 0.8 in my analysis . Figure ?? (Appendix) depicts graphically how the probability changes at the cutoff.

In equation (2), outcome variable Y is at the individual level and β_1 is the coefficient of interest. I also include state fixed effects (δ_s) and cluster the standard errors at the district level. The estimated β_1 is the treatment effect and can be interpreted as a local average treatment effect (LATE). To estimate this effect, I apply non parametric methods using data around the cutoff within a chosen bandwidth and using a triangular kernel with a linear specification. As a result, the estimates obtained are based on data closer to the cutoff as opposed to the entire sample. The optimal bandwidth size is based on Calonico et al. (2019) which minimizes the mean square error of the RD point estimator. Besides using the optimal bandwidth for my primary results, I check for robustness to other bandwidths, uniform kernel and polynomial of degree two.

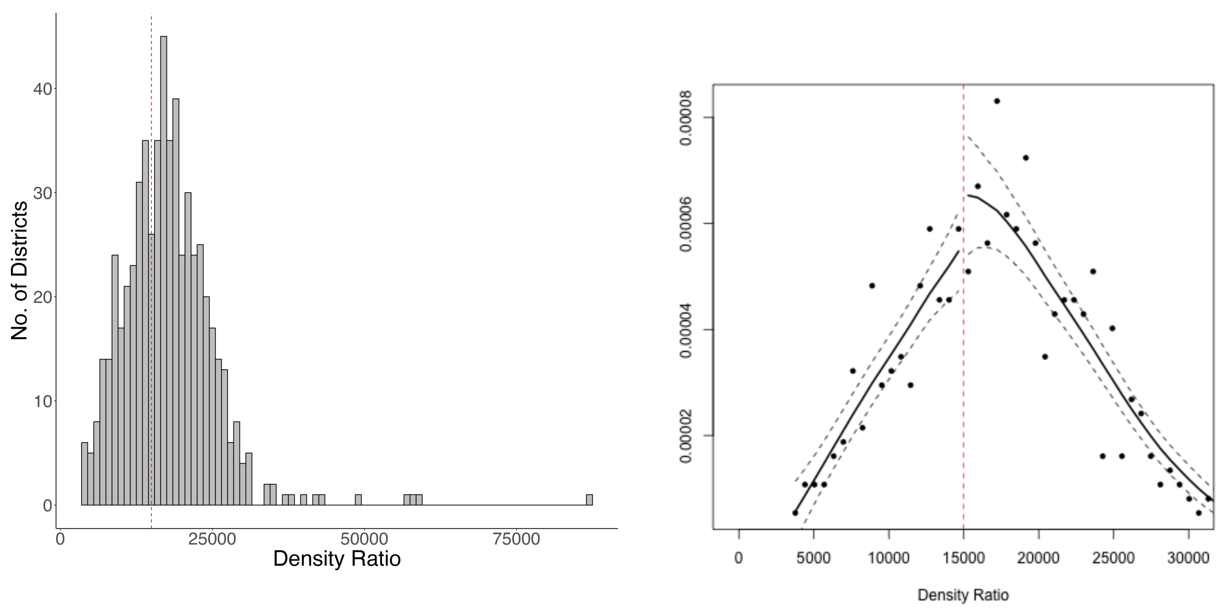
4.1 Identification

There are two conditions that hold in this policy setup which allow me to use regression discontinuity. First is the presence of a continuous eligibility measure or a running variable in the form of density ratio. In this setup, I am able to rank the districts based on this density ratio measure. Second, is the presence of a clearly defined threshold in the form of the national average of density ratio. This gives us a point above which the districts are eligible to be classified as underbanked. However, these conditions should be accompanied by validity of two underlying assumptions in order to interpret the treatment effect as causal.

First, I need to ensure that there was no manipulation of density ratio around the cutoff. If some districts wanted to take advantage of the underbanked status, they could try to manipulate their density ratio such that it barely exceeds the threshold, making them eligible for treatment. In such a scenario we would see a cluster of districts with density ratio slightly above the threshold. For this to be false, the density of running variable should be smooth around the cutoff. In other words, the number of districts with density ratio just above the threshold and just below the threshold needs to be similar. To test this assumption, I run the McCrary (2008) density test which gives a p-value of 0.35. This suggests that I should not reject the null hypothesis of density being smooth around cut off. Figure 1 shows a finely-gridded histogram of running variable and a smoothened histogram using local linear regression, separately on either side of the cut-off.

Figure 1: Distribution of Density Ratio

(a) Histogram **(b)** McCrary Density Test



(a) shows the distribution of density ratio around the cutoff and (b) shows the McCrary density test with a p-value = 0.35. The dashed red line represents the cutoff = 14976.

Another reason to believe no manipulation around cutoff could be the method of constructing the running variable. The numerator of density ratio uses population data from 2001 which couldn't have been manipulated unless districts anticipated this policy four years in advance. The denominator uses the total number of bank branches existing in 2005. It is unlikely that the districts could manipulate this number since it includes different types of banks, operated by private as well public sector enterprises.

Manipulation of the running variable is directly associated with another important assumption of treatment and control districts around the cutoff being similar in the absence of this policy. Smoothness of pre-determined characteristics around the cutoff is an assumption that lies at the heart of identification in an RD setup. This could be invalidated if treatment assignment is public knowledge, which holds true over here. If some districts could manipulate their density ratio to become a part of the treatment group and these districts happen to be inherently different from others, then our assumption would fail. Hence, it is important to ensure pre-policy balance such that all variables except for treatment are smooth around the cutoff before the policy. To see this, I show the control and treatment mean in Table 1 before the policy was implemented. I also run the fuzzy RD with pre-policy variable data.

Table 1: Pre-Policy Balance

	Non Underbank		Underbank		RD Coef
	Mean	SD	Mean	SD	
Banking Variables : RBI					
New Bank Openings (Pvt)	1.92	4.75	0.31	0.81	0.047 (0.131)
New Bank Openings (Govt)	3.63	5.71	1.01	2.23	0.162 (0.215)
Total Banks (Pvt)	26.69	50.24	4.03	10.33	-0.062 (0.180)
Education Variables : ASER					
Above 13 Enrollment (Girls)	0.84	0.36	0.81	0.39	-0.014 (0.027)
Below 13 Enrollment (Girls)	0.89	0.32	0.88	0.32	-0.011 (0.025)
Above 13 Enrollment (Boys)	0.86	0.35	0.84	0.37	-0.031 (0.026)
Below 13 Enrollment (Boys)	0.89	0.31	0.89	0.31	-0.008 (0.022)
Individual Variables : ASER					
Age	10.24	3.17	10.02	3.18	0.186* (0.108)
Never Enrolled	0.01	0.12	0.02	0.15	0.004 (0.006)
Dropout	0.03	0.17	0.04	0.19	0.025** (0.009)
Private Schooling	0.25	0.43	0.18	0.38	0.04 (0.025)
Household Variables : ASER					
Household Size	5.88	3.43	6.58	3.79	0.156 (0.277)
Mother's Schooling	0.55	0.50	0.40	0.49	-0.08 (0.05)
Mother's Age	31.81	9.53	32.30	9.51	0.049 (0.703)
Household Banking Variables : IHDS					
Debt	0.37	0.48	0.47	0.5	-0.026 (0.10)
Source : Banks	0.36	0.48	0.22	0.41	0.008 (0.07)
Source : Moneylender	0.25	0.43	0.36	0.48	-0.05 (0.10)
Source : Relatives	0.14	0.35	0.21	0.41	0.09* (0.05)
Purpose : Marriage	0.14	0.37	0.16	0.37	-0.007 (0.04)
Purpose : Education	0.04	0.19	0.02	0.12	0.009 (0.02)
Employment Variables (16-29 yrs) : NSS					
LFP	0.56	0.5	0.54	0.5	-0.016 (0.028)
Employed	0.50	0.50	0.49	0.50	0.00 (0.026)
LFP(Lower Caste)	0.59	0.49	0.55	0.49	-0.034 (0.027)
Employed(Lower Caste)	0.54	0.49	0.52	0.49	-0.014 (0.029)
MPCE(log)	8.08	0.68	7.89	0.64	-0.017 (0.046)

5 Results

In this section, I document the effects of the bank expansion policy on underbanked districts. I start with the first stage outcomes i.e. whether there was any impact on banking growth in the districts that were classified as underbanked according to the policy. This helps ascertain the effectiveness of this policy on direct outcomes that were being targeted. Next, I look at the impact on educational outcomes, specifically enrollment across both genders and different age groups. The aim is to identify whether banking growth led to any changes in a household's decision to invest in education, and to see if there is any treatment effect heterogeneity based on age group and gender. To support the findings on enrollment, I examine if households are increasing their expenditure on education. In addition, I also explore utilization of banking services at the household level to shed some light on whether the households were credit constrained and how they made use of the increased access to financial services.Lastly, I look at the labor market outcomes to see if certain groups benefitted more than others.

5.1 Effects on growth of private sector banks

The primary analysis in reference to this bank expansion policy would be to first see whether there was any response from the banks. Did the policy incentivize banks to target the underbanked districts for new branches? Since the incentives were mainly targeted towards private banks, I focus on the private banking growth here.

As anticipated, I find significant effects for private banks as opposed to no effect on public banks. Comparing underbanked and non-underbanked districts seven years after the policy was introduced, I find significant discontinuity in the number of new private bank branches that were established. This shows us that by 2012, the number of new private banks established in treatment districts was significantly higher than the control districts. Table 2 depicts the regression discontinuity coefficient on treated districts when looking at the new bank branch openings until 2012. Seven years after the policy, treated

Table 2: New Banks Until 2012

	Private	Public
Treated	0.279**	-0.015
	(0.133)	(0.118)
Control Mean	2.37	3.43
First Stage	0.81	0.81
Bandwidth	3483	3253
N Eff.	225	217
N	565	571
Controls	Yes	Yes

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Variables have been winsorized at 95 percent level.

(underbanked) districts have around 27 percent more private bank openings than control districts . However, comparing the new public bank openings, I find that treatment districts were as likely to get new government banks as control districts. The coefficient for public banks is very small and insignificant. Hence, over and above the usual growth in bank infrastructure in India, the added incentives to private banks through this policy led to their higher growth in underbanked districts. These regression results are accompanied by a graphical analysis using an RD plot that highlights the discontinuity around cutoff. (Appendix Figure 6)

Next, I look at how the dynamics of bank expansion varied across different years following the introduction of this policy. For this, I look at the new private bank branch openings and total number of private bank branches in each year (Figure 2).After one year of this policy, I see that private banks respond by opening around 35 percent more bank branches on average in underbanked districts. This treatment effect seems to be particularly large and statistically significant at 1 percent level in 2009 , after which it turns insignificant. This could be attributed to changes in the underlying list of underbanked districts introduced by RBI in 2010[4]. I also draw comparisons in the level of private sec-

[4] RBI introduced changes that allowed banks to open branches without obtaining any licenses in some

Table 3: Total Private Banks

	Mar 2006	2006	2007	2008	2009	2010	2011	2012	2013	2014
Treated	-0.062	-0.019	0.006	0.067	0.216	0.310**	0.250*	0.105	0.125	0.135
	(0.180)	(0.187)	(0.166)	(0.145)	(0.144)	(0.147)	(0.147)	(0.139)	(0.150)	(0.155)
Control Mean	2.23	2.26	2.4	2.55	2.57	2.66	2.85	2.98	3.21	3.3
First Stage	0.81	0.81	0.81	0.81	0.82	0.82	0.81	0.8	0.8	0.8
Bandwidth	3146	3227	3450	3241	3795	3684	3142	2829	2831	2676
N Eff.	209	216	223	216	243	236	209	192	192	181
N	568	568	568	568	568	568	568	568	568	568
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

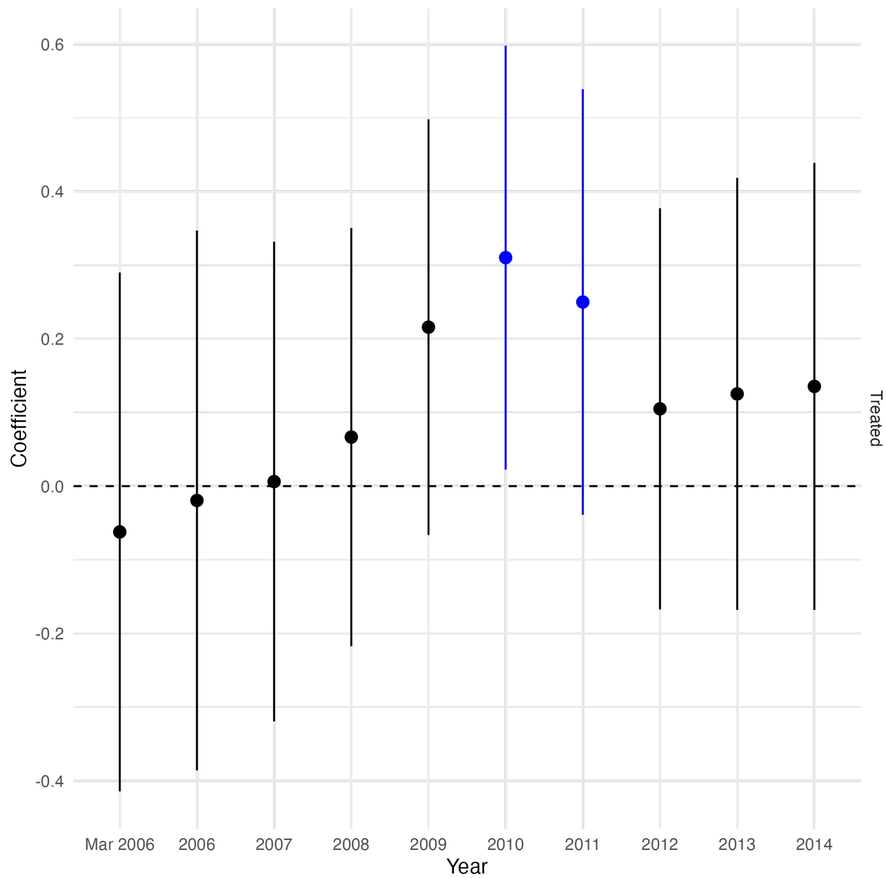
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Variables have been winsorized at 95 percent level.

tor banking across districts by examining the total number of private banks. Referring to Table 3, we can see a gradual increase in the treatment effect post policy, with the treatment districts having around 36 percent more private banks than control districts in 2010. The increased growth in banking in underbanked districts until 2009 led to a significantly higher number of private banks, which lasted even after the growth slowed down.

The effects detailed above can be interpreted as causal only under the assumption that the pre-treatment banking variables vary smoothly around the cutoff. We want to ensure that pre-policy, there is no discontinuity in new private bank openings and the number of private banks at the cutoff used for the expansion policy. This is evident when we look at the treatment coefficient in 2005 for both the variables, which is small and insignificant. Hence, we could interpret these results as the causal effect of bank expansion policy on establishment of private banks in underbanked districts.

of the districts falling in control group.

Figure 2: Total Number of Private Bank Branches



Summarizing the first stage results, I find that the RBI policy incentivized private sector banks to expand their operations in the districts classified as underbanked. There was a higher presence of brick-and-mortar private banks in these districts in comparison to others. Seeing no effect on public sector banks post policy and establishing smoothness of banking variables pre-policy provides evidence in support of this expansion being driven by the underlying policy. Now, I'd like to examine if financial development in underbanked districts also resulted in better educational outcomes.

5.2 Effects on educational outcomes

Given that the policy had a positive effect on availability of private banks in underbanked districts, I'd like to see whether that led to changes in educational outcomes as well. Tying the results from previous literature, we see that this policy led to higher economic development, changes in the job market and boosted business activity. These developments on a macro level are also likely to have an impact on households, thereby affecting decision making with regard to investing in their child's education. To study the impact on education, I use school enrollment data from the ASER 2016 household level survey. The survey records whether a child within the age of 5-16 years is enrolled in school or not. Since I see a prominent first stage impact around 2012, it is likely that the spillovers to the education sector are lagged, because of which I look at outcomes from 2016.

5.2.1 Effects on education status for girls

I first consider the entire sample of individuals between 5-16 yrs to examine if the children in treatment districts are more likely to be enrolled in school in comparison to control. Table 4 specifies the treatment effect from running an RD around the cutoff with an MSE-optimal bandwidth chosen by Calonico et al.'s method. Dependent variable denotes the likelihood of being enrolled in school. While there is no effect on boys' enrollment, bank expansion leads to a positive and significant effect on girls' school enrollment. Girls in underbanked districts are 3.4 percentage points more likely to be in school than girls in non-underbanked districts. Given a control mean of 0.95, this corresponds to a 3.5 percent increase in overall enrollment of girls. This shows that an increased presence of brick-and-mortar banks in districts improves the educational status of females, making them more likely to enroll and stay in school.

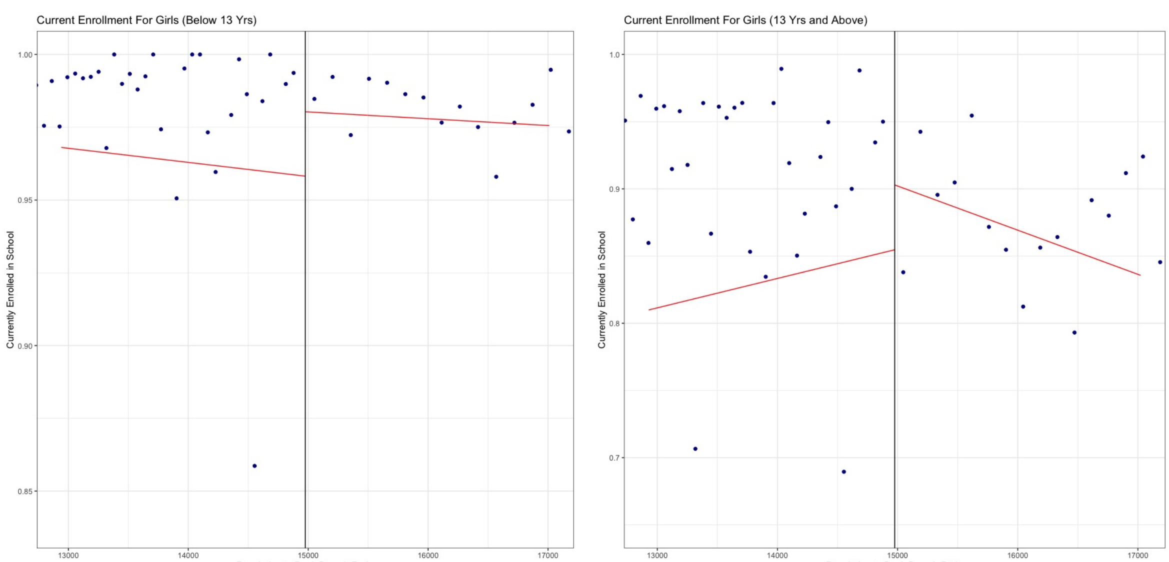
Table 4: Current Enrollment (2016)

	All Ages		Below 13 Yrs		Above 13 Yrs	
	Girls	Boys	Girls	Boys	Girls	Boys
Treated	0.034*	0.006	0.026*	0.005	0.056**	0.010
	(0.017)	(0.008)	(0.014)	(0.005)	(0.028)	(0.016)
Control Mean	0.95	0.96	0.98	0.98	0.91	0.92
First Stage	0.87	0.86	0.87	0.87	0.86	0.85
Bandwidth	1939	2638	2035	2842	2047	2963
N Eff.	40810	63782	29396	46620	14745	24070
N	217043	227859	143577	153481	73466	74378
Controls	Yes	Yes	Yes	Yes	Yes	Yes

* p < 0.1, ** p < 0.05, *** p < 0.01.Variables have been winsorized at 95 percent level.

Figure 3: RD Plots (ASER 2016)

(a) Enrollment (Girls Below 13 yrs) **(b)** Enrollment (Girls Above 13 yrs)



Educational attainment of females in developing nations such as India can be a factor of financial constraints as well as prevailing social and cultural norms. These constraints can be even stronger when we look at low income or historically disadvantaged groups. Existing norms around traditional gender roles can further lead to under-investment in adolescent girls' education. Hence, to understand how the underlying costs associated with female education evolve with age, I examine the treatment effects for individuals specifically in their teenage years. As girls enter their teenage years, costs associated with secondary education can be higher. This could be attributed to increasing financial burden and prevalence of early marriage in Indian context.

Restricting my sample to individuals above the age of 12, I find that the treatment effect is larger and positive for adolescent girls and still insignificant for adolescent boys. Adolescent girls in treatment districts are 5.6 percentage points more likely to be enrolled in school in comparison to girls in the same age group in control districts. I can conclude that the bank expansion policy leads to a 6 percent increase in probability of being in school in underbanked districts for adolescent girls. While I still see a 2.6 percent increase amongst younger girls, having a very high control mean makes it less likely to tease out an effect in that group. Figure 3 shows the discontinuity graphically in an RD plot.

Contextually, adolescent girls in India are the ones at margin and face a higher risk of dropping out than adolescent boys. While the gender gap in primary education has reduced significantly, the gap in secondary education still remains. For a household, decision to invest in a child's education could be determined by the degree of financial constraints as well as the opportunity cost associated with schooling. While the financial constraints could be dictated by the household economic status or caste , the opportunity cost can vary significantly based on the child's gender. Coupled with age, being an adolescent can further increase the cultural barriers as well as opportunity cost for girls. This includes participation in household chores, marriage, menstruation or employment opportunities. Hence, who is at the margin is determined by an interaction of age as well as gender. To further elaborate on how the dynamics change over different age groups for both genders, I look at the treatment effect heterogeneity by age.

Table 5: Current Enrollment by Age Group and Gender

Age (Yrs)	5-6	7-8	9-10	11-12	13-14	15-16
Girls						
Treated	0.005	0.018	0.035**	0.033*	0.046*	0.066*
	(0.009)	(0.015)	(0.014)	(0.018)	(0.027)	(0.034)
Control Mean	0.99	0.98	0.98	0.97	0.94	0.86
First Stage	0.87	0.88	0.86	0.85	0.87	0.86
Bandwidth	3904	1948	2376	2158	2056	2013
N Eff.	8657	7654	10710	9094	8055	6660
Boys						
Treated	0.000	0.004	0.009*	0.002	0.004	0.032
	(0.006)	(0.005)	(0.005)	(0.009)	(0.009)	(0.024)
Control Mean	0.99	0.99	0.99	0.98	0.96	0.89
First Stage	0.86	0.88	0.88	0.85	0.87	0.86
Bandwidth	2298	3894	2666	2560	4063	3785
N Eff.	5432	16851	12381	12082	16960	13691

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Variables have been winsorized at 95 percent level.

5.2.2 Do the effects differ by age group?

In the second step of my analysis, I disentangle the effects of this policy by age and demonstrate that the positive effects on enrollment of girls increase as we go up the age group. For boys, I continue to find insignificant and small effects for all age groups, except 9-10 yrs old.

Table 5 depicts the RD coefficient if I split my sample by different age groups. I find an increase in probability of school enrollment of girls in underbanked districts that begins around the age of 9 and corresponds a 3 to 8 percent increase over the control mean. For girls in the age group 11-12 years, I observe a 3.4 percent increase in the likelihood of enrollment over the average in control districts. This treatment effect is the highest among girls in the age group of 15-16 years, at 7.6 percent increase over the control group.

In summary, the bank expansion policy led to an improvement in enrollment status of girls in underbanked districts and the effect is even more pronounced among higher age groups. To interpret these results as causal, I would have to ensure that there aren't any pre-policy discontinuities in the enrollment variable around the cutoff. To do so, I consider ASER 2007 data which is the earliest possible round available. The RD treatment coefficients obtained for different age groups of girls are all insignificant. This provides evidence in support of the assumption that pre-treatment enrollment for girls varied smoothly around the cutoff. Figure ??(Appendix) draws a comparison of RD coefficient for different age groups of girls pre-policy and post-policy.

The results that I observe here could be driven by two plausible reasons. Having a high enrollment rate in control districts for younger girls as well as boys could be contributing to little to no treatment effects for younger girls. Second, I would expect contextual mechanisms to be at play in the background that are contributing to this heterogeneity. The educational opportunities that girls have access to in India are based on multiple factors, such as economic and cultural, which make it more difficult for them to continue secondary education. Hence, I would anticipate any positive effects I find, to be larger when girls are at a higher risk of dropping out. In another section, I explore the potential channels through which bank expansion maybe affecting these underlying factors.

6 Robustness Checks

To show that my results are robust, I use multiple robustness checks in a regression discontinuity setup. The first one is to check sensitivity of my results to multiple bandwidths around the cutoff (density ratio = 14976). In my main results, I use the MSE optimal bandwidth based on Calonico et al's approach. I test for robustness to bandwidths falling within a range of 0.5 to 2 times the optimal bandwidth. Table 6 shows the RD coefficient with varying bandwidths for enrollment. For girls, I find that the results remain qualitatively similar and significant for all but one alternate bandwidth. I do not find any treatment effect for boys' enrollment irrespective of the bandwidth chosen. Similar

Table 6: Sensitivity to Different Bandwidths (ASER Enrollment)

	0.5	0.75	1	1.25	1.5	1.75	2
Girls							
Treated	-0.010	0.024	0.034*	0.032*	0.028*	0.025*	0.024*
	(0.022)	(0.018)	(0.017)	(0.016)	(0.015)	(0.014)	(0.013)
Bandwidth	969	1,454	1,939	2,424	2,908	3,393	3,878
N.Eff.	20,455	31,684	40,810	56,668	67,575	77,646	86,847
Boys							
Treated	0.003	0.003	0.006	0.007	0.008	0.009	0.009
	(0.009)	(0.009)	(0.008)	(0.008)	(0.007)	(0.007)	(0.006)
Bandwidth	1,319	1,978	2,638	3,297	3,956	4,616	5,275
N.Eff.	30,943	45,588	63,782	80,541	92,096	108,231	119,912

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

Table 7: RDD With Uniform Kernel (ASER Enrollment)

	Below 13 Years		Above 13 Years	
	Girls	Boys	Girls	Boys
Treated	0.024*	0.005	0.059**	0.009
	(0.014)	(0.006)	(0.029)	(0.017)
Bandwidth	2,053	1,402	1,785	2,373
N.Eff.	29,611	22,019	12,714	19,107

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

analysis for subsample of different age groups of girls is included in the appendix.

Next, I test for robustness to a uniform kernel (Table 7) which assigns equal weight to all observations within the optimal bandwidth. Standard rdrobust method with MSE optimal bandwidth uses triangular kernel which attaches a higher weight to the observations closer to cutoff. Even with the uniform kernel, my results remain robust.

The third robustness check I implement is to run a donut regression discontinuity (Table 8). This involves running the same RD after removing data points around the treatment cutoff to address any systematic sorting of districts near the cutoff. I consider

Table 8: Donut RD (ASER Girls' Enrollment)

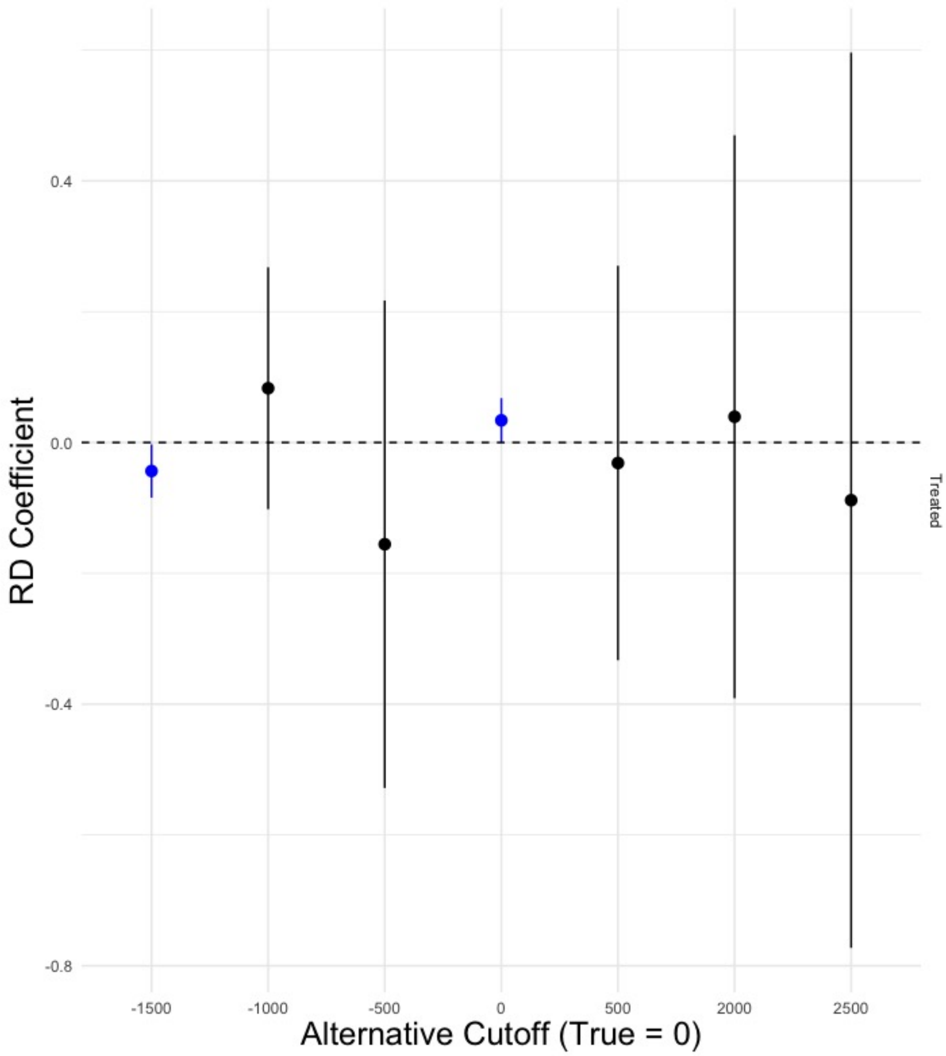
	All Observations	Excluding (-250,+250) Cutoff	Excluding (-500,+500) Cutoff
Treated	0.034*	0.066***	0.021*
	(0.017)	(0.023)	(0.012)
Bandwidth	1,939	2,378	3,498
N.Eff.	40,810	51,590	70,654
N	217,043	212,746	207,732

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

two bandwidths, +-250 and +-500 around the cutoff to exclude the observations. I still notice a significant positive effect on girls' enrollment after excluding observations around the cutoff.

Another test that I run involves checking for a discontinuity around placebo cutoffs (Figure 4). For this, I consider multiple cutoffs in the range of +-3000 of the treatment cutoff such that there is enough data around the placebo cutoff to run an RD. I find my primary outcome i.e. girls' enrollment to be smooth around all but one placebo cutoff of -1500 (density ratio = 13476). Upon examining the data around this level of density ratio, I find the data to be very volatile such that districts very close and to the right of this placebo cutoff have low enrollment whereas districts to the left have high enrollment. Since my primary regression uses triangular kernel within a bandwidth of approximately 2000 and attaches a lower weight to observations far away from the cut-off, the data points around -1500 should not impact my results significantly. Lastly, I test for a different functional form using a polynomial of order two and find the results to be qualitatively similar (Appendix).

Figure 4: Robustness to Placebo Cutoff



7 Potential Channels

Having established that girls' education improves, the next question would be how does a higher availability of banks in an area lead to a higher investment in girls' education? To explore this link between financial inclusion and schooling, I investigate three potential channels. First is increased savings through a higher access to financial products. The other two stem from the general equilibrium effects of this bank expansion policy, in the

Table 9: Household Debt and Investment Variables

		Investment		
	Debt	Bank Account	Microfinance account	Savings (Post Office)
Treated	-0.110	0.072	0.002	0.187***
	(0.088)	(0.087)	(0.033)	(0.064)
Control_Mean	0.58	0.5	0.09	0.22
First_Stage	0.78	0.78	0.75	0.77
Bandwidth	1,787	1,972	2,913	1,707
N.Effective.	9,385	10,364	15,786	8,889
N	40,081	39,968	39,958	39,947
Controls	Yes	Yes	Yes	Yes

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.Variables have been winsorized at 95 percent level.

form of economic growth leading to higher income and better labor market outcomes for households.

7.1 Are the households credit constrained?

The primary objective of this policy was to incentivize banks to make inroads into areas that were previously lacking access to banking infrastructure. While the first stage results show that the policy was successful in achieving this, I now turn to examining how households adopted these banking services. There could be two ways in which households could benefit from increased banking, thereby affecting schooling outcomes. First, they can access formal credit more easily, either for education loan or for other purposes, which frees up resources and provides them financial stability. Second could be access to a savings account which serves as a tool to accumulate assets as well as generate income.

Table 9 depicts the results using IHDS (2012) household level data. I find that this policy doesn't lead to an increase in borrowing in underbanked districts. Qualitatively, it indicates that the likelihood of taking a debt in the last 5 years was 19 percent lower in underbanked districts. However, I find that the probability of saving money in a post office account or scheme increases by 18.7 percentage points. Given a low control mean of 0.22, this corresponds to an 85 percent increase on average over the control districts. In other words, if every fourth individual participated in some form of savings with the post office in non-underbanked districts, then this would be true for every second individual in underbanked districts. The probability of having a bank account is also higher but insignificant. This suggests that easier availability of credit doesn't lead to higher borrowing. Rather, presence of financial services in these districts leads households to save their money in an account.

Apart from analysing borrowing on the extensive margin in terms of a change in the likelihood of taking a loan, I also investigate the intensive margin of borrowing to see if there was any substitution of informal sources of credit with the formal sources of credit. In Table 10, I find that individuals are 12.8 percentage points or 50 percent less likely to borrow from a microfinance institution in underbanked districts. Borrowing from a bank increases by 13.5 percent even though it is statistically insignificant. This substitution across various sources of credit could be particularly true for households belonging to low income or lower caste groups as they face high interest rates and discrimination in the informal market. When restricting my sample by caste, there is some evidence that lower caste households are more likely to borrow from their employer in underbanked districts. I elaborate further on the outcomes for lower caste households later in this section. Conditional on taking a loan, I also find that households are more likely to take a loan for educational purposes and less likely for marriage expenses in underbanked districts relative to non-underbanked.

Summarizing, I see strong and significant effects on savings amongst households which could help them generate income, inculcate savings behavior as well as smooth consumption in the long run. This could be a likely channel contributing to better educational outcomes for girls.

Table 10: Credit Variables

	Source					Purpose	
	Bank/Credit Program	Micro-finance	Money lender	Employer	Friends/ Relatives	Marriage expenses	Educational
Treated	0.053	-0.128***	-0.094	0.011	-0.099	-0.037*	0.039***
	(0.063)	(0.043)	(0.063)	(0.018)	(0.063)	(0.022)	(0.014)
Control Mean	0.39	0.27	0.31	0.05	0.45	0.17	0.05
FS	0.82	0.82	0.84	0.85	0.85	0.86	0.86
BW	2,112	2,132	1,718	1,654	1,618	1,594	1,566
N Eff.	6,103	6,122	5,051	4,930	4,899	4,709	4,672
N	22,014	22,014	22,014	22,014	22,014	21,959	21,959
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes

* p < 0.1, ** p < 0.05, *** p < 0.01. Variables have been winsorized at 95 percent level.

7.2 Are households spending more towards education?

To further look into the income effect channel, I investigate how the household expenditure on education is changing. In alignment with the results related to higher economic growth and business activity, it is likely that an increase in income and expenditure could be leading to higher retention in schools. For this, I look at education expenditure data from NSS 2014 Education survey. While I find suggestive evidence of household expenditure on books increasing in the underbanked districts by around 12 percent, the effect is not significant (p-value = 0.312) . Comparing the expenditure in level terms, I see that the results are qualitatively similar and still insignificant. The expenditure results here capture the effect of individuals who are moving from no schooling to schooling with low expenditure and the ones who were possibly affected only in terms of the amount they spent on education.

To further investigate the income channel, I turn to general equilibrium effects of this policy on labor market outcomes. These effects could vary when we look at low income households or the households belonging to lower caste who have been historically poorer.

Table 11: Labor Market and Savings

	LFP	Employed	Weekly Wages (log)	Saving Bank
Lower Caste(SC/ST/OBC)				
Treated	0.049**	0.042*	0.135**	0.062**
	(0.023)	(0.023)	(0.066)	(0.028)
Control Mean	0.48	0.44	6.81	0.19
First Stage	0.88	0.88	0.89	0.87
Bandwidth	3,339	3,797	2,054	2,417
N Eff.	36,171	39,891	6,092	20,824
All Castes				
Treated	0.039*	0.035	0.195**	0.014
	(0.023)	(0.024)	(0.082)	(0.026)
Control Mean	0.45	0.42	6.87	0.15
First Stage	0.82	0.81	0.86	0.83
Bandwidth	3,594	3,277	1943	3,772
N Eff.	52,756	49,552	5,154	45,759

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.Variables have been winsorized at 95 percent level.

Therefore, I investigate the outcomes for a sample of households belonging to the lower caste separately in the next section.

7.3 Labor Market Outcomes for Lower Caste

Here, I specifically focus on individuals from household's belonging to SC, ST or OBC to see if the positive effects of this policy on education are driven by the households at the bottom of the distribution. Households belonging to these castes have historically been poorer, lack education and are often discriminated against which makes them more likely to turn to informal sources of credit. These households are also the ones where the gender gap in education is higher, which could be attributed to a combination of traditional norms and economic disadvantage.

In Table 11, I depict the RD coefficient for labor market outcomes of a subsample belonging to lower castes and in the age group 16-29 yrs. This suggests that lower castes in underbanked districts were 10 percent more likely to be participating in labor force and 10 percent more likely to be employed in comparison to lower caste individuals in non-underbanked districts. Their weekly wages are also increasing by 2 percent and the likelihood of having access to a savings bank increases by 32.6 percent. Improvement in labor market outcomes as well access to banks could be linked to better schooling outcomes for these households, specifically for girls who are at a higher risk.

In general, I see an improvement in labor force participation as well as wages for the entire population. However, the wages here are available only for non self-employed individuals which doesn't allow me to account for the shift in type of jobs available as well movement of individuals across different sectors.

Summarizing the potential channels, I find higher savings and some suggestive evidence of increased education spending by households in underbanked districts. Combining this with the general equilibrium effects of this policy, I see that individuals benefit from higher employment as well higher wages, specifically for individuals at the bottom of the distribution which is proxied by caste. An interaction of who benefits the most from this policy in terms of income and who is at the margin in schooling could be a reason why I observe positive effects only for girls and not boys.

8 Conclusion

In this paper I look at the impact of increased access to banking services on girls' education. In 2005, Reserve Bank of India categorized districts as underbanked and non-underbanked based on an arbitrary cutoff and incentivized private banks to penetrate those areas. Using this as a source of exogenous variation in the availability of brick-and-mortar banks, I show that the policy leads to an improvement in schooling outcomes for girls in terms of enrollment. Exploiting branch-level data of banking network in India, I show that there was significantly higher private banking growth in underbanked districts. Using household level data on schooling in India, I show that girls in underbanked districts were more likely to be in school in comparison to girls in non-underbanked districts. There is treatment effect heterogeneity, with older girls benefiting the most. In contrast, I do not find any effects for boys. I provide sugestive evidence on increased savings and income that could be leading to these effects. These results are likely to be driven by economically and historically disadvantages groups, where the gender gap in schooling is high.

References

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

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

Ashraf, Nava; Dean Karlan; and Wesley Yin. 2010. "Female Empowerment: Impact of a Commitment Savings Product in the Philippines," World Development 38(3): 333–344. https://www.sciencedirect.com/science/ article/abs/pii/S0305750X09001910.

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., E. Breza, E. Duflo, and C. Kinnan (2019). Can microfinance unlock a poverty trap for some entrepreneurs? Technical report, National Bureau of Economic Research.

Barboni, Giorgia, Erica Field, and Rohini Pande, "Rural Banks Can Reduce Poverty:

Evidence from 870 Indian Villages," 2021.

Breza, Emily Kinnan, Cynthia , 2021. "Measuring the Equilibrium Impacts of Credit: Evidence from the Indian Microfinance Crisis," The Quarterly Journal of Economics, President and Fellows of Harvard College, vol. 136(3), pages 1447-1497.

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

Brune, Lasse and Gine, Xavier and Goldberg, Jessica and Yang, Dean and Yang, Dean, Commitments to Save: A Field Experiment in Rural Malawi (August 1, 2011). World Bank Policy Research Working Paper No. 5748, Available at SSRN: https://ssrn.com/abstract=1904244

Burgess, Robin and Rohini Pande, "Do rural banks matter? Evidence from the Indian social banking experiment," American Economic Review, 2005, 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.

Chandar, Bharat and Jiao, Dian and Mo, Marshall, "Bank Expansion, Firm Dynamics, and Structural Transformation: Evidence from India's Policy Experiment", (May 1, 2023). Available at SSRN

Calonico, Sebastian, MatiasDCattaneo, and MaxHFarrell, "Optimal bandwidth choice for robust bias-corrected inference in regression discontinuity designs",The Econometrics Journal, 2020, 23 (2), 192–210

Chiapa, C., Prina, S., and Parker, A. (2016) The Effects of Financial Inclusion on Children's Schooling, and Parental Aspirations and Expectations. J. Int. Dev., 28: 683–696. doi: 10.1002/jid.3137.

Cramer, Kim Fe, "Financial Development and Health", Available at SSRN 3917526,2021.

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

KABOSKI, J. P. AND R. M. TOWNSEND(2011): "A Structural Evaluation of a Large-Scale Quasi-Experimental Microfinance Initiative," Econometrica, 79, 1357–1406. ———(2012): "The impact of credit on village economies," American economic journal. Applied economics, 4, 98.

Koomson, Isaac Afoakwah, Clifford. (2022). Can financial inclusion improve children's learning outcomes and late school enrolment in a developing country?. Applied Economics. 55. 10.1080/00036846.2022.2086683.

Iqbal, Kazi Roy, Paritosh K. Alam, Shamsul, 2020. "The impact of banking services on poverty: Evidence from sub-district level for Bangladesh," Journal of Asian Economics, Elsevier, vol. 66(C).

Mahajan Kanika, Ritadhi S.K., Kulkarni Nirupama, "Banking the Underbanked: Capital Investment and Credit-Constrained Firms"

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

Prina, Silvia, 2015. "Banking the poor via savings accounts: Evidence from a field experiment," Journal of Development Economics, Elsevier, vol. 115(C), pages 16-31.

Parker, Susan Vogl, Tom "Do Conditional Cash Transfers Improve Economic Outcomes in the Next Generation? Evidence from Mexico" The Economic Journal, Volume 133, Issue 655, October 2023, Pages 2775–2806, https://doi.org/10.1093/ej/uead049

Raghuram G. Rajan Luigi Zingales , 1996."Financial Dependence and Growth", NBER Working Papers 5758, National Bureau of Economic Research, Inc.

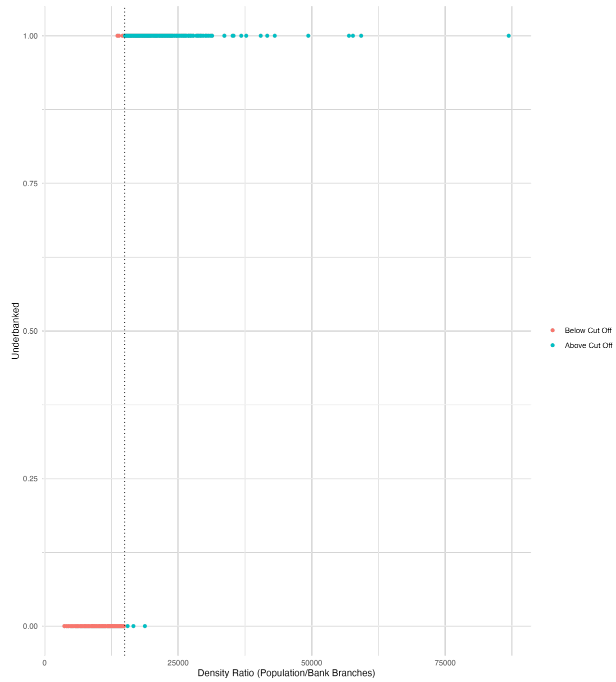
Susan W Parker, Tom Vogl, Do Conditional Cash Transfers Improve Economic Outcomes in the Next Generation? Evidence from Mexico, The Economic Journal, Volume 133, Issue 655, October 2023, Pages 2775–2806, https://doi.org/10.1093/ej/uead049

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

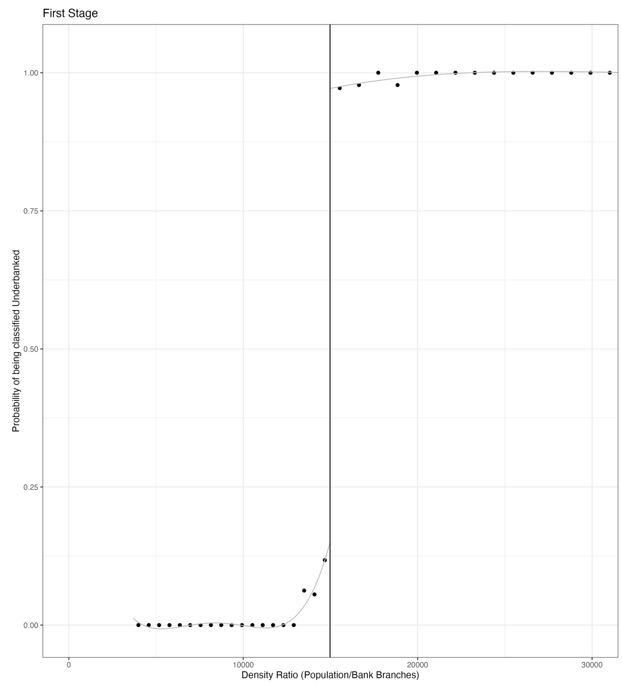
Appendix

Figure 5: First Stage

(a) Fuzzy Setup



(b) First Stage Regression



Note : (a) shows the fuzzy RD design and (b) shows the first stage regression results graphically.

Table 12: Summary Statistics : ASER 2016

	All Obs.		Obs. in Band-width	
	Non-underbank	Underbank	Non-underbank	Underbank
Density Ratio	11181.92 (2740.16)	22192.5 (7362.47)	13728.74 (789.74)	16313.72 (825.43)
Individual				
Child Age	10.85 (3.12)	10.76 (3.15)	10.78 (3.12)	10.78 (3.15)
School Class	5.31 (3.12)	4.93 (3.12)	5.09 (3.11)	5.1 (3.1)
Never Enrolled	0.01 (0.09)	0.02 (0.12)	0.01 (0.1)	0.01 (0.1)
Dropout	0.02 (0.15)	0.04 (0.2)	0.03 (0.17)	0.04 (0.19)
Tution	1.7 (0.58)	1.62 (0.63)	1.67 (0.6)	1.66 (0.62)
Tution Amount	50.15 (175.65)	56.37 (177.2)	50.34 (166.38)	49.28 (182.43)
Govt. School	0.6 (0.49)	0.66 (0.47)	0.6 (0.49)	0.67 (0.47)
Pvt. School	0.36 (0.48)	0.28 (0.45)	0.36 (0.48)	0.29 (0.45)
Household Level				
HH Type	2.37 (0.79)	1.98 (0.88)	2.26 (0.82)	2.05 (0.88)
Electricity	1.06 (0.26)	1.22 (0.43)	1.09 (0.31)	1.12 (0.34)
Toilet	1.27 (0.46)	1.53 (0.52)	1.36 (0.49)	1.44 (0.51)
TV	1.22 (0.43)	1.53 (0.52)	1.31 (0.48)	1.37 (0.5)
Motor Vehicle	1.55 (0.52)	1.67 (0.49)	1.59 (0.51)	1.59 (0.52)
Mobile	1.13 (0.37)	1.21 (0.45)	1.17 (0.42)	1.21 (0.45)
Father's Schooling	1.1 (0.5)	1.22 (0.55)	1.16 (0.52)	1.19 (0.53)
Mother's Schooling	1.28 (0.5)	1.49 (0.55)	1.35 (0.53)	1.43 (0.55)
Village Level				
Pucca Road	0.84 (0.36)	0.82 (0.39)	0.83 (0.38)	0.83 (0.37)
Post Office	0.48 (0.5)	0.35 (0.48)	0.45 (0.5)	0.41 (0.49)
Bank	0.35 (0.48)	0.26 (0.44)	0.31 (0.46)	0.27 (0.45)
Primary School	0.9 (0.3)	0.93 (0.25)	0.9 (0.31)	0.93 (0.25)
Private School	0.45 (0.5)	0.39 (0.49)	0.48 (0.5)	0.38 (0.49)

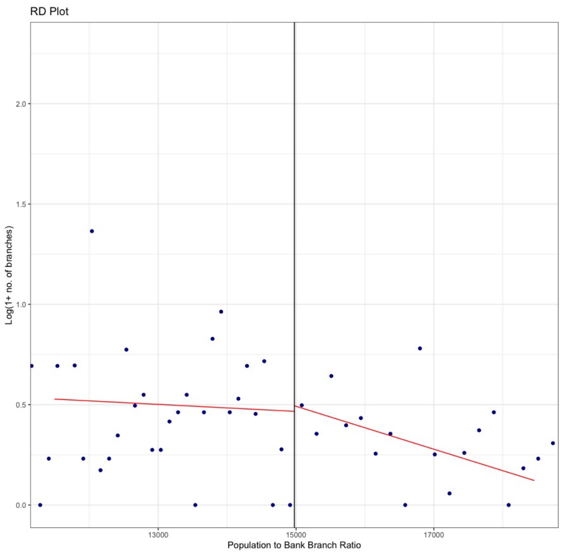
I consider a bandwidth of [-2500,2500] for summary statistics. The table displays variable mean and standard deviation (in parenthesis).

Table 13: New Private Banks

	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
Treated	0.047	-0.303	0.300*	0.214	0.720***	-0.107	-0.268	-0.218	0.037	-0.071
	(0.131)	(0.200)	(0.154)	(0.168)	(0.271)	(0.153)	(0.164)	(0.252)	(0.217)	(0.165)
Control Mean	0.5	0.65	0.81	0.65	0.78	0.89	1.38	1.5	1.68	1.41
First Stage	0.81	0.79	0.8	0.81	0.8	0.8	0.81	0.8	0.81	0.8
Bandwidth	3482	2039	2747	3069	2449	2419	2934	2496	3430	2490
N.Eff.	225	129	185	204	166	162	198	170	222	169
N	565	565	565	565	565	565	565	565	565	565
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.Variables have been winsorized at 95 percent level.

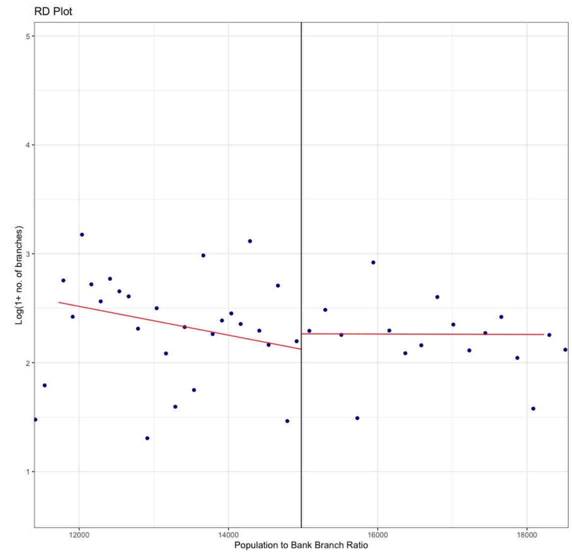
Figure 6: Banking Outcomes : RD Plot



(a) Pre-policy (2005)

X-axis: Density Ratio

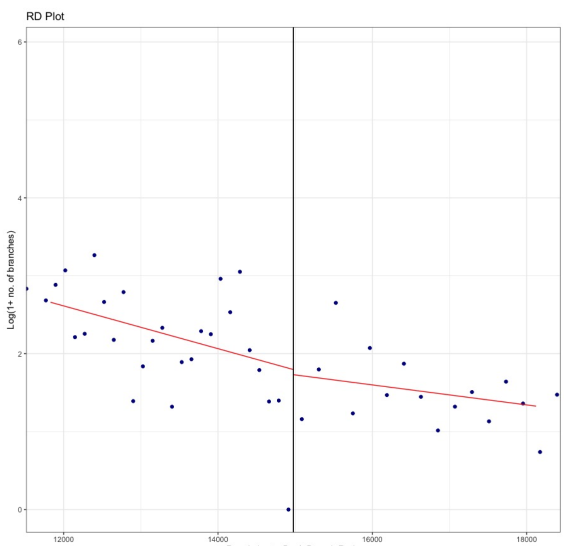
Y-axis: New Private Banks Opened



(b) Post Policy (Until 2012)

X-axis: Density Ratio

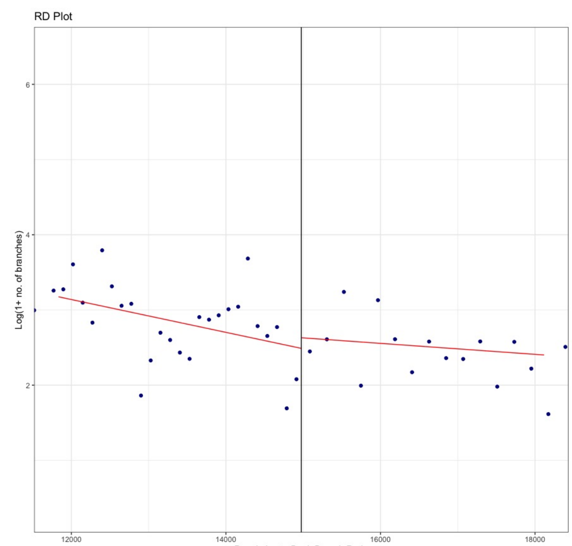
Y-axis: New Private Banks Opened



(c) Pre-Policy (Mar 2006)

X-axis: Density Ratio

Y-axis: Total Private Banks



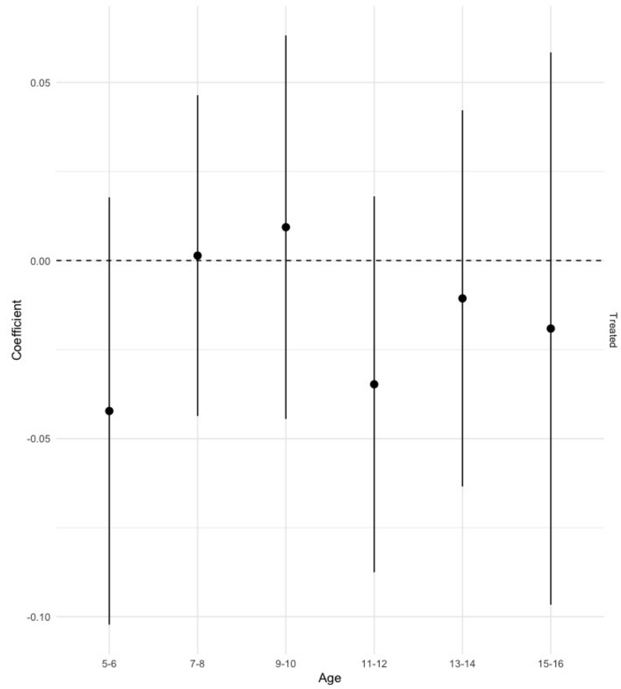
(d) Post Policy (2011)

X-axis: Density Ratio

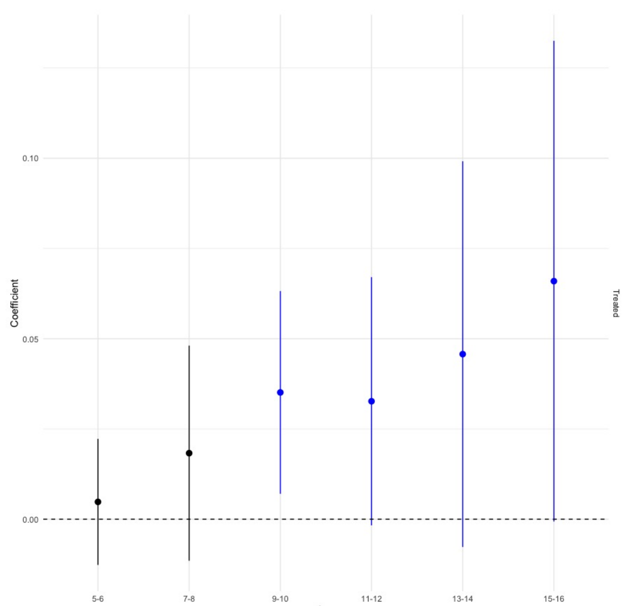
Y-axis: Total Private Banks

Figure 7: RD Coefficient for Girls Enrollment

(a) Pre-policy (2007)



(b) Post-policy (2016)



Here, X-axis denotes the age groups (5-6 yrs; 7-8 yrs; 9-10 yrs; 11-12 yrs; 13-14 yrs; 15-16 yrs) and Y-axis denotes the RD treatment coefficient obtained for different age group samples.

Table 14: Enrollment by Age Group and Gender (Pre-policy 2007)

Age (Yrs)	5-6	7-8	9-10	11-12	13-14	15-16
Girls						
Treated	-0.042	0.001	0.009	-0.035	-0.011	-0.019
	(0.031)	(0.023)	(0.027)	(0.027)	(0.027)	(0.040)
Control Mean	0.73	0.93	0.94	0.93	0.88	0.74
First Stage	0.84	0.83	0.77	0.79	0.81	0.81
Bandwidth	4,230	3,561	2,706	2,678	2,341	2,973
N.Effective.	20,303	22,088	18,056	16,499	11,942	10,201
Boys						
Treated	-0.014	0.010	-0.013	0.006	-0.007	-0.081**
	(0.033)	(0.022)	(0.025)	(0.025)	(0.027)	(0.033)
Control Mean	0.73	0.94	0.94	0.94	0.9	0.8
First Stage	0.82	0.78	0.77	0.79	0.81	0.82
Bandwidth	2,594	2,435	2,850	2,819	2,936	2,801
N.Effective.	15,029	17,709	22,351	19,822	17,568	11,965

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Variables have been winsorized at 95 percent level.

Table 15: Sensitivity to Different Bandwidths (ASER Enrollment)

	0.5	0.75	1	1.25	1.5	1.75	2
Girls (Below 13 Years)							
Treated	0.002	0.020	0.026*	0.027**	0.025**	0.023**	0.022**
	(0.015)	(0.014)	(0.014)	(0.013)	(0.012)	(0.012)	(0.011)
Bandwidth	1,018	1,526	2,035	2,544	3,053	3,562	4,070
N.Effective.	14,789	21,481	29,396	39,391	47,055	53,507	60,382
Girls (Above 13 Years)							
Treated	0.001	0.047	0.056*	0.050*	0.038	0.034	0.034
	(0.039)	(0.031)	(0.029)	(0.026)	(0.024)	(0.023)	(0.022)
Bandwidth	1,023	1,535	2,047	2,559	3,070	3,582	4,094
N.Effective.	7,450	10,815	14,745	19,809	23,975	27,115	30,937

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

Table 16: RD With Polynomial Degree 2 (ASER Enrollment)

	Below 13 Years		Above 13 Years	
	Girls	Boys	Girls	Boys
Treated	0.024	0.004	0.036	-0.014
	(0.016)	(0.006)	(0.033)	(0.021)
Bandwidth	3,668	3,911	3,817	3,524
N.Eff.	54,633	61,266	29,176	27,734

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

Table 17: Summary of Placebo MSE-Optimal Cutoff Results

Alternative MSE-Optimal Cutoff	**Bandwidth**	**RD Coef**	**p-value**	**CI Lower**	**CI Upper**	**Obs. (Left)**	**Obs. (Right)**
-1500	802.75	-0.04	0.04	-0.08	-0.00	7666	7999
-1000	698.79	0.08	0.38	-0.10	0.27	5801	8401
-500	1318.92	-0.16	0.41	-0.53	0.22	13831	4222
0	1938.90	0.03	0.05	-0.00	0.07	19133	21677
500	2557.75	-0.03	0.84	-0.33	0.27	5089	36228
1000	1339.44	-2.41	0.45	-8.63	3.81	10754	20412
1500	1264.89	-0.28	0.20	-0.70	0.15	14919	19978
2000	1303.67	0.04	0.86	-0.39	0.47	17189	20407
2500	1028.79	-0.09	0.80	-0.77	0.60	16624	14063