Impact of Education Loans on Higher Education: The Indian Experience*

Senjuti Patra Tridip Ray Arka Roy Chaudhuri
Indian Statistical Institute

December, 2017

Preliminary draft, please do not cite

Abstract

Escalating costs of higher education globally have made the effectiveness of credit access in increasing educational attainment an important subject for both academics and policy makers. This paper presents evidence of a strong positive impact of education loan availability on tertiary education outcomes in India. We evaluate the performance of a model education loan scheme introduced by the Indian central bank in 2001 that has drastically increased the availability of education loans in India. First, we estimate the impact of the increased availability of education loans on the years of schooling attained by an individual by exploiting the variation in the number of education loan accounts across districts and the exposure to the program across age cohorts. We find that one standard deviation increase in the number of education loan accounts leads to a 6.17% improvement in years of schooling over the sample mean. Then we estimate the effect of availability of education loans on an individual's decision to enrol for higher education, by making use of across district and over time variation in the number of education loan accounts. The results suggest that one standard deviation increase in the number of education loan accounts results in a 6.87% increase in higher education enrolment over the sample mean. We also find heterogeneous effects of education loans on enrolment and years of schooling, with the effects being more pronounced for the relatively disadvantaged groups across gender, caste and location (rural/urban).

*The paper benefited greatly from discussions with Farzana Afridi, S. Chandrashekhar, Deepti Goel, Abhiroop Mukhopadhyay, Bharat Ramaswami, Ranjan Ray, Renuka Sane and seminar participants at the Eighteenth World Congress of the International Economic Association (IEA), Mexico City; IEA-IGIDR Conference on Social Sector Development in India, IGIDR, Mumbai; International Seminar on Innovations in Financing of Higher Education, NUEPA, New Delhi and seminar participants at Ashoka University and Indian Statistical Institute.

1 Introduction

Higher Education is a pivotal driving force for economic growth, and of paramount importance in economic and social mobility. Vast research literature including papers by Hanushek and Kimko (2000); Krueger and Lindahl (2000); Hanushek and Woessmann (2007) provides evidence of the value of investing in education to develop human capital and of its contribution to economic development and growth. Recent studies [1] in a diverse set of countries show that the returns to post-secondary education has been increasing compared to the returns to primary education. Consequently, making higher education available and affordable to the masses occupies the center stage in academic and policy discussions in all countries, developed and developing alike.

In light of the increasing demand for and escalating costs[2] of higher education, higher education to the masses is not an easy plan to implement. In this context, education loan is an important and arguably effective tool for financing higher education. Education loans have been the order of the day in many developed countries for quite some time now. However, in India, education loans have gained impetus as a mode of financing higher education only in recent years, particularly after the introduction of a model education loan scheme in 2001 launched by the Reserve Bank of India (the central bank of India) in consultation with the Indian Banks' Association. In this paper we study the effect of this recent improvement in availability of education loans on individuals' higher education decisions.

We evaluate the impact of availability of education loans separately on years of schooling and on enrolment in tertiary education. We use across district and across age cohort variation in exposure to the improved availability of education loans to identify the effect of education loans on years of schooling. The effect of education loans on enrolment is identified by making

[1] See, for example, Schultz (2003); de Ferranti et al. (2003); Manacorda, Sanchez-Paramo, and Schady (2005) and Kingdon and Soderbom (2007a, 2007b)

[2] According to the report of the 71^{st} round of the NSS Education survey that we discuss in the data section, the average annual private expenditure on general education has increased by 176% and that on technical and professional education has increased by 96% between 2008 and 2014.

use of the across district and over time variation in loan availability. Both exercises generate statistically significant positive effects of education loans. We also identify heterogeneity in effects of loan availability on tertiary eduction outcomes across social and demographic groups, with the effects being larger in general for the historically disadvantaged sections.

The Indian economy having leapfrogged into a service led growth regime since the 1990s, the need for catering higher education to the masses has become imperative to sustain the momentum of growth. It is through investment in human capital that the country can effectively exploit the demographic dividend generated by the high concentration of young adults in the population. In India, family income or economic status has always been an important determinant of college attendance. From the National Sample Survey (NSS) data[3] that we use for this study, we see that the participation in higher education has been largely skewed in favour of the highest consumption expenditure[4] quartile. Well designed education loans can make higher education affordable by helping households overcome credit constraints and reap the benefits of the substantial returns to higher education. If adequately inclusive, education loans can also help poor households to come out of low income-low skill equilibrium traps.[5] Despite the potential importance of education loans in improving higher education enrolment and educational attainment, there has been no empirical research on the impact of education loans on higher education in India. In a study that is first of its kind, we estimate the effects of education loan availability on tertiary education outcomes.

While there has been extensive research on the effect of education loans on various outcomes including enrolmet and persistence in higher education, course choice and occupa-

[3]For a detailed description please refer to the data section of the paper.

[4]Consumption expenditure is a reasonable indicator of economic status in absence of reliable data on income.

[5]The link between imperfect credit access and poverty traps has been extensively studied in a number of papers, including those by Galore and Zeira (1993) and Piketty (1997). Evidence of the importance of credit constraints in education decisions is discussed in Attanasio and Kaufmann (2009), Belley and Lochner (2007), Lochner and Monge-Naranjo (2011a) and Brown, Scholz and Seshadri (2012). Specific empirical evidence in favour of education loans increasing participation of low income families has been provided in Solis(2012). We come back to this paper later in this section.

tional choice, majority of these papers are in the context of developed countries.[6] Given significant differences in returns to tertiary education and the demand for and supply of skills across developed and developing countries,[7] sparsity of the literature evaluating the impact of education loans in developing country contexts is a major lacuna. This paper presents robust estimates of the impact of education loan availability on higher education outcomes in a developing country. Unlike most papers in the United States that use educational institution level policy changes to identify the effect of education loans, we use a nationally representative sample to evaluate the impact of a country-wide shift in education loan availability. By looking at both enrolment in tertiary education and years of schooling, we capture effects of education loans not only on participation in tertiary education but also on overall educational attainment.

Our paper is related to the broader literature investigating the relevance of credit constraints in education. Access to loans can be expected to make a difference in enrolment or educational attainment if credit constraints are important for education decisions. However, the evidence on the importance of borrowing constraints in education is mixed. Using data from the US to analyse enrolment in higher education, studies like Cameron and Heckman (1998), Cameron and Heckman (2001), Carneiro and Heckman (2002) and Cameron and Taber (2004) attribute the difference in college enrolment rates across the rich and the poor to differences in ability generated by systematically lower investment in the early stages of education by poor households. On the other hand, recent studies by Belley and Lochner (2007), Lochner and Monge-Naranjo (2011a) and Brown, Scholz and Seshadri (2012) argue that borrowing constraints are a major deterrent to participation in higher education for individuals form low income families. Using data on mothers' and students' expectations of returns to schooling from Mexico, Attanasio and Kaufmann (2009) provide evidence of the

[6]See, for example, Dunlop (2012), Wiederspan (2015) and Rothenstein and Rouse (2010). These papers have been discussed later in this section.

[7]See Montenegro and Patrinos (2013); Sianesi and Van Reenen (2000) for an overview of the variation in returns to education across low and high income countries.

important role of credit constraints in enrolment decisions.

The estimation of the effect of loans on educational outcomes is often difficult because students self select their loan amount. Researchers try to employ quasi-experimental approaches by exploiting an exogenous variation in access to loans. Using variation in community colleges' participation in the federal loan program in the US, Dunlop (2012) finds no effect of access to federal loans on college attendance. However, she finds negative effects of limiting loan access on degree attainment and the transfer to four year colleges. Wiederspan (2015) also analyses the effect of community colleges' participation in the loan program. However, unlike Dunlop who uses only cross-sectional variation across institutions, Wiederspan uses within institution, across time variation to filter out institution specific factors. He finds no effects on attendance, degree attainment and transfer to a four year college while identifying a positive effect of borrowing on the number of credit courses. Rothenstein and Rouse (2010) analyse the effect of student debt on early career choices by making use of a similar policy shift in a reputed American university. They find that student debt significantly lowers graduates' probability of choosing low paid, public interest jobs. Solis (2012) uses two education loan programs in Chile that had loan eligibility perfectly determined by test scores to employ a regression discontinuity design to estimate the causal impact of access to credit on enrolment. He finds positive and statistically significant effects. Solis also identifies a heterogeneity in the effects with respect to quartiles of family income. The effect is strongest for the poorest quartile and weakest for the richest quartile. In a similar study from a developing country, Gurgand, Lorenceau and Mélonio (2011) use credit score cutoffs to evaluate a loan program in South Africa. They find similar positive effects.

We evaluate the impact of loan availability on tertiary education outcomes of all individuals who could have benefited from the introduction of the education loan scheme, based on their age and district of residence. In this intent-to-treat analysis framework, the concern about self selection is mitigated. Our results indicate positive and statistically significant effect of

the education loan program on enrolment, similar to Solis (2012) and Gurgand, Lorenceau and Mélonio (2011). However, unlike Solis (2012), we identify the strongest positive effects of loan availability on enrolment for the 2^{nd} and 3^{rd} quartiles of land holding[8]. The effects are statistically insignificant for the poorest and the richest land quartiles.

This paper estimates the effects of education loan availability on enrolment and educational attainment in the context of tertiary education in India. Since household or individual level information on amount borrowed for education is unavailable, we use the total number of education loan accounts in a district as an indicator of the extent of availability of education loans in the district. To identify the effect of the education loan program on years of schooling, we utilise the fact that the exposure to the education loan scheme varies across districts and age cohorts. We compare a cohort that completed higher education before the introduction of the education loan scheme in 2001 with another cohort that enroled in higher education after the introduction of the scheme and thus had full exposure to the program. In estimating the impact of the education loan program on enrolment, we exploit the variation in the number of education loan accounts across districts and over time. We include district fixed effects, time fixed effects and some relevant district and time varying variables to rule out possible sources of endogeneity. We find positive, statistically significant and robust effects of education loan on both years of schooling and enrolment in tertiary education. We also find heterogeneous effects of education loans on years of schooling and enrolment, with the magnitude of the effect being larger in general for disadvantaged groups.

The rest of the paper is organized as follows. Section 2 gives a brief overview of the Education Loan Scheme. The datasets that we use are described in section 3. Sections 4 and 5 study the impact of education loans on years of schooling and enrolment in tertiary education respectively. Each of these sections is further divided into subsections discussing identification strategy and results. We conclude in section 6.

[8] A proxy for income in rural areas.

2 The Education Loan Scheme

In 2001, the Government of India in consultation with the Reserve Bank of India (RBI) and the Indian Banks' Association (IBA) laid down the guidelines of the model Education Loan Scheme. The main objective of the Educational Loan Scheme is to provide financial support to meritorious but poor students, through the banking system, to pursue higher education in India and abroad. The banks were advised by RBI to implement the scheme. In addition to this, education loans were made part of priority sector lending[9]. Most banks have formulated their own versions of an education loan scheme based on the IBA guidelines. We next discuss some features of the model education loan scheme.

Any Indian National who has secured admission in recognized institutions in India and abroad is eligible. The scheme covers graduation and post-graduation as well as diploma courses that are approved by UGC or AICTE[10]. Apart from tuition fees, expenditure on travel, examinations, stationery, books, etc is also considered for the loan.

A student can borrow up to INR 1,000,000 for studies in India and upto INR 2,000,000 for studying abroad. Margins and collateral requirements vary across loan amounts. Interest rates vary across banks. The repayment starts 6 months after getting a job or 1 year after completion of course, whichever is earlier. Repayment is made in fixed Easy Monthly Installments (EMI). The loan must be repaid within 10 years (for loans up to INR 750,000) or 15 years (above INR 750,000).

In 2009-10, the government announced an interest subsidy scheme for students from economically disadvantaged sections(family income less than INR 450,000 a year) pursuing technical

[9]Under priority sector lending regulations, the Reserve Bank of India mandates banks to provide a specified portion of the bank lending to few specific sectors like agriculture and education that are instrumental for economic development and might not receive adequate credit in absence of the regulation.

[10]University Grants Commission (UGC) is a statutory body under the Ministry of Human Resource Development, Government of India, that provides recognition to universities. All India Council for Technical Education (AICTE) is an autonomous statutory institution under the UGC that supervises technical and management education in India.

or professional studies in India. Under this scheme, the interest accrued during the moratorium period will be paid by the government. To make the loans further attractive, income tax rebates are also offered on the amounts of repayment of an education loan.

The number of loan accounts under this scheme has grown steadily over the years, with the major chunk of it coming from nationalised banks. The contribution by regional rural banks, though small in magnitude, has also grown steadily. We use this scheme to evaluate the performance of education loans in the tertiary education scene in India.

3 Data

In this paper, we evaluate the effect of improved education loan availability separately on educational attainment or the years of schooling achieved by individuals, and on contemporaneous enrolment in tertiary education. For both the exercises, we use district level data on the total number of education loan accounts to construct the main independent variable of interest. This information has been provided by Indian central bank, the Reserve Bank of India in it's dataset for Basic Statistical Returns of Scheduled Commercial Banks and is available for the years 2001 to 2015. The dataset divides the total number of active education loan accounts in each year according to bank group - public sector banks, private sector banks and regional rural banks and by population group - rural and urban.

We use the 71^{st} round of the NSS[11] education survey conducted in 2014 for evaluating the impact of education loans on years of schooling. The NSS dataset contains information on several household and individual level characteristics like age composition of the household, religion, caste[12], land holding, highest level of education achieved, enrolment status and so

[11]NSS data refers to data from surveys conducted by the National Sample Survey Organisation (NSSO), an organisation under the Ministry of Statistics and Program Implementation, Government of India that conducts regular nationally representative socio-economic household surveys in India.

[12]Indian society is marked by the existence of a system of social stratification by caste. The 'general' category refers to individuals belonging to upper castes. The 'Scheduled Castes' or SCs refer to historically disadvantaged lower castes. The 'Scheduled Tribes' or STs are the indigenous people of India. 'Other Backward Castes' or OBCs refer to the collective of castes that have been socially disadvantaged.

on. We use the information on the highest level of education achieved to construct our dependent variable, years of schooling[13]. In addition to the demographic controls sourced from the NSS dataset, we also control for time varying district level variables that might possibly confound the results of our regression. We use information from the All India Survey of Higher Education, conducted in 2014, on the district and year wise number of colleges and universities to control for the total number of higher education institutions in a district. We also control for the number of bank branches in a district using year wise information on open and closed bank branches in each district from the Reserve Bank of India's Directory of Commercial Bank Offices.

For estimating the effect of education loan availability on contemporaneous enrolment, we use individual and household level information from three rounds of the NSS Employment & Unemployment Surveys: 61^{st} round conducted in 2004-05, 66^{th} round conducted in 2009-10 and 68^{th} round conducted in 2011-12. We control for number of bank branches and number of higher education institutions by making use of the previously mentioned data sources. In addition to bank branches and higher education institutions, we also control for night light intensity per square kilometer area of a district, a commonly used proxy for GDP.[14] This information is sourced from the DMSP[15] night lights database.

[13]We map education levels to years of education according to the standard duration of completion of each of the education levels in India.

[14]See Henderson et al. (2012), Pinkovskiy (2013), Michalopoulos and Papaioannou (2013).

[15]Defense Meteorological Satellite Program Operational Linescan System archives data on nighttime light intensity starting from 1992. Information on the specific location of night lights is sourced from the National Geophysical Data Center (NGDC).

4 Effect of Education Loan Availability on Years of Schooling

4.1 Identification Strategy

In this section we investigate the impact of better loan availability on overall educational attainment, the years of schooling achieved by individuals. For all the analysis that follows, we need to define a higher education age group. Individual and household level data from four rounds of the NSS surveys: the 61^{st}, 66^{th} and 68^{th} rounds of the employment & unemployment surveys and the 71^{st} round of the education survey show that enrolment in higher education (graduate and above) is concentrated in the age group 17-25 years (Refer to figure 1.), with more than 90% of the individuals enroled in higher education in each year belonging to this age group. Henceforth, we will be referring to this age group as the higher education age group.

The most recent dataset with detailed information on educational attainment of individuals is the 71^{st} round of the NSS education survey conducted in 2014. Since we are interested in the effect of improved loan availability on the final educational attainment, we should only be considering individuals who were old enough to have completed higher education by 2014. According to our definition of the higher education age group, this would mean looking at individuals at least 26 years of age in 2014.

Of all individuals older than 26 years of age in 2014 in a particular district, different age cohorts had different exposures to the education loan availability according to the time when they were in the higher education age group. A cohort that completed tertiary education before 2001 (that is, was 25 years of age before 2001) did not benefit at all from the introduction of the education loan scheme and the consequent accessibility of education loans. All younger cohorts had differential exposure to the improved availability of education loans, depending on the district of residence and the time when they were enrolled in tertiary ed-

ucation. We make use of this between cohort and across district variation in exposure to identify the effect of improved education loan availability on the total years of schooling achieved by an individual. We employ a difference in difference strategy similar to Duflo (2001) and compare a cohort with non-zero exposure to the program to a cohort that has completed higher education by 2001 with zero exposure to the program.

We assign treatment and control status to cohorts based on the time when they were in the higher education age group. We keep the sizes of both cohorts the same as that of the higher education age group. The treatment cohort is the youngest cohort that completed higher education before 2014, meaning that the youngest individuals in the treatment cohort have to be 26 in 2014. Considering the size of the cohort and the age of the youngest individual, our treatment cohort is the cohort consisting of 26-34 year olds in 2014. The control cohort is the youngest cohort with zero exposure to the treatment- the youngest person in the control cohort has to be at least 25 by 2001. Given the age of the youngest member and the size of the cohorts, the control cohort consists of 38-46 year olds in 2014.[16]

We next need a measure of the extent of exposure of the treatment group to the education loan program. The number of education loan accounts in an individual's district of residence at the time when she is in the higher education age cohort is indicative of the extent of her exposure to the improved availability of education loans. The treatment cohort in this exercise, 26-34 year olds in 2014, was the in higher education age cohort, or was in the 17-25 years age group, in 2005. We thus take the number of education loan accounts in 2005 in the district of residence as a measure of the treatment group's exposure to the education loan program.

We allow for the educational attainment of the treatment cohort to be different from that of the control cohort by including a cohort fixed effect. We control for district level, cohort

[16]Since there is no strict age limit for availing education loans, there is a possibility that the treatment cohort has also benefited to some extent from the improved availability of education loans post 2001. However, this would imply that our estimates are conservative, and we are underestimating the positive impact of education loans on educational attainment.

invariant factors that might be correlated with both educational attainment of individuals and the number of education loan accounts by taking district fixed effects. The regression specification is as follows:

$$Y_{idc} = \alpha + \beta(T_i * L_d) + \gamma \mathbf{X}_{idc} + \eta_d + \mu_c + \epsilon_{idc} \tag{1}$$

Y_{idc}, the outcome variable of interest, is the years of schooling of individual i living in district d belonging to cohort c. L_d is the number of education loan accounts created in district d in 2005 (per hundred population in 2005) and T_i is a dummy variable indicating whether individual i belongs to the younger (or treatment) cohort. The interaction term, $T_i * L_d$, determines program intensity, i.e., the extent to which an individual has been exposed to the treatment. The coefficient β on this interaction term is our main coefficient of interest. $\mathbf{X}_{idc}$ is a vector of individual-level controls such as sex, caste, sector, etc. η_d represents district fixed effects and μ_c represents the cohort fixed effects. As stated earlier, we have two cohorts: the younger or treatment cohort (26-34 year olds in 2014) and the older control cohort (38-46 year olds in 2014). Table 1 presents the summary statistics for the main independent variable and the dependent variable.

4.2 Results

4.2.1 Main Regression

The results of the regression on years of schooling (regression specification described in Equation (1)) have been outlined in Table 2. The coefficient of interest, which is the coefficient on the interaction between the treatment dummy and number of education loan accounts per hundred population in district of residence in 2005, is positive and statistically significant. The number of years of schooling of the treatment cohort (26-34 year olds in 2014) increased by 3.75 years for each unit increase in the number of loan accounts per hundred population in 2005. This signifies a 6.17% increase in years of education of the treatment cohort per

standard deviation increase in the measure of loan availability. We thus capture statistically significant positive effects of the education loan program on educational attainment of individuals.

4.2.2 Heterogeneous Effects on Years of Schooling

The existing disparities in access to education are evidenced by the coefficients of some of the demographic controls in the main regression on years of schooling. Number of years of schooling achieved is lower for individuals residing in rural areas and consistently higher for males. Educational attainment is also significantly lower for members of the scheduled castes, scheduled tribes and other backward castes as compared to members of the upper castes. It would be of interest from a policy perspective to investigate the differential impacts of the education loan program on such disadvantaged groups. In order to identify and understand the heterogeneity in the effect of improved education loan availability on educational attainment in different social and demographic groups, we divide the data and run the main regression specification on each of theses groups. In the rest of this section, we discuss heterogeneous effects of education loans on years of schooling across gender, caste groups and sectors.

Table 3 presents the results of the main regression with the sample divided according to gender. The coefficient of interest is statistically significant and positive for females and insignificant for males. The difference between the coefficients from the two subpopulations is also significantly different from zero. The program thus seems to have had a greater positive effect in increasing the years of schooling of females.

We have then divided the sample according to castes. Table 4 reports the results of the regression in these subsamples. The effect of the program is positive and statistically significant for scheduled castes and scheduled tribes and other backward castes and insignificant for other upper castes. The differences between the coefficients of SC/ST and upper castes,

and that of the OBC and upper castes are statistically significant . That is, as before, we find a greater positive impact of the program on the disadvantaged social groups.

In Table 5, we present the results separately for the rural and urban samples. Both coefficients are positive and statistically significant, but the magnitude of the coefficient is higher for the rural sample. This difference between the magnitudes of the coefficients is statistically significant. That is, while education loan availability has a significant positive effect on the years of schooling of individuals in both rural and urban areas, the magnitude of the effect is higher in rural areas.

4.3 Falsification Exercise

The coefficient of interest from the main regression might be vulnerable to bias if the trend in increase in educational attainment varies systematically across districts. Our results will be misleading if the districts where educational attainment was increasing faster were also the districts that saw a greater expansion of education loans. However, such trends should be captured in systematic differences in educational attainment among older cohorts that were not exposed to the program as well. We make use of the presence of multiple cohorts with zero exposure to the program in our data to test this implication. We consider two cohorts with zero exposure to the treatment: 38-46 year olds in 2014 and 50-58 year olds in 2014. The younger of the two is the youngest cohort with zero exposure to the treatment, the control group from the main regression specification. This is our new "treatment" cohort. The older cohort, which has been chosen keeping age difference between the two cohorts and the size of the cohorts same as those in the main regression, becomes our control cohort. Using these definitions of treatment and control groups, we run the regression specification in Equation (1). If the results of the main regression are not being driven by pre-existing trends, this regression should not capture any effects of the program on the "treatment" group's years of schooling. The results of this regression have been presented in Table 6. The coefficient of interest is not statistically significant for the control regression. This helps

in validating our claim about the results not being driven by pre-existing district specific trends.

4.4 Robustness Checks

The main regression so far does not control for district level cohort varying factors that might be correlated with educational attainment and our independent variable of interest. The district level numbers of educational institutions and bank branches are two obvious candidates for possible confounders. Districts that have seen a greater rise in the number of higher education institutions since 2001 might also be likely to have a greater number of education loan accounts in 2005. The effect of education loans that we are identifying will then be conflated with the effect of increase in number of educational institutions on educational attainment. Systematic correlation between increase in bank branches and the creation of education loan accounts might also bias our results through a similar channel.

In order to demonstrate that our results are robust to the inclusion of educational institutions and bank branches, we control for bank branches and institutions in our main regression in various ways. The results have been presented in Table 7. In the second column, we check if the pre-treatment values of these variables are systematically related to the program and overall educational attainment. We control for the interaction of the cohort dummy and the pre-treatment (corresponding to the year 1999) values of these variables in the regression specification. In column 3, we control for the number of educational institutions and bank branches in an individual's district of residence at the time of her entry into the higher education age group (that is, when she was 17 years old). The specification in column 4 controls for the nine year average number of educational institutions and bank branches while an individual was in the higher education age cohort (was 17-25 years of age). The statistically significant positive effect of the education loan program, as measured by the coefficient of program intensity, is preserved across all specifications.

In another set of robustness checks presented in Table 8, we control for age and household fixed effects.By controlling for age fixed effects, we are controlling for an overall trend in educational attainment that is common to all districts. In the regression controlling for household fixed effects, we are comparing individuals from the same household but different cohorts. By using household fixed effects, we are limiting the set of possible confounders to household level cohort variant factors, which are unlikely to be syatematically correlated with the district level intensity of the education loan program. The coefficient of program intensity remains positive and statistically significant for all plausible combinations of cohort, district, age and household fixed effects.

4.5 Effect of Education Loans on the Highest Level of Education Achieved

While the effect on years of education gives us an idea about the effect of improved loan availability on overall educational attainment, we are also interested in the differential impact of loan availability at different levels of education. Breaking down the impact of education loans at different levels of education might shed some light on the mechanism through which the increase in educational attainment is being effected. Given substantial returns to higher education[17], households might be willing to invest in the earlier stages of education only if there is a possibility of participation in college education in the future. Apart from the obvious channel of increasing years of schooling by making higher education affordable and making it easier for students to finish college, the loan program might also be reducing dropouts at earlier stages of education by making credit constrained households more optimistic about the possibility of enrolment in higher education in the future. We run the following regression to estimate the effect of education loans on the probability of a level of education m (primary, secondary, higher secondary, graduate and post graduate) being the

[17]Recent studies show that the returns to higher education has been increasing compared to the returns from primary education. See, for example, Kingdon and Soderbom (2007a, 2007b).India-specific evidence can be found in Kingdon and Unni (2001).

highest level of education an individual achieves:

$$Y_{idcm} = \alpha + \beta(T_i * L_d) + \gamma \mathbf{X}_{idc} + \eta_d + \mu_c + \epsilon_{idc} \tag{2}$$

Y_{idcm}, the dependent variable, takes value one if the highest level of education attained by individual i living in district d and belonging to cohort c is m. L_d, as in the previous specification, is the number of education loan accounts in district d in 2005 (per hundred population in 2005) and T_i is a dummy variable indicating whether individual i belongs to the younger (or treatment) cohort. The interaction term, $T_i * L_d$, determines program intensity. The coefficient β on this interaction term is our main coefficient of interest. $\mathbf{X}_{idc}$ is a vector of individual-level controls. η_d represents district fixed effects and μ_c represents the cohort fixed effects. As stated earlier, we have two cohorts: the younger or treatment cohort (26-34 year olds in 2014) and the older control cohort (38-46 year olds in 2014).

The results of this regression have been presented in Table 9. There is a statistically significant positive impact of the program at the graduate and postgraduate levels of education, with the magnitude of the effect being greater for graduates. This indicates that the program increases the likelihood of completing college education. There is a negative impact of the program on the probability of primary education being the highest level of education achieved by an individual. This might be reflecting an increase in the number of people who complete primary education and attain higher levels of education. However, in 2005, the program is still relatively new and it is too early to substantiate any claims about the program reducing drop outs at lower levels of education by increasing the expectations about the probability of enroling for higher education in the future.

5 Effect of Education Loan Availability on Enrolment in Higher Education

5.1 Identification Strategy

So far, we have identified positive and significant effects of education loans on years of schooling and attainment of graduate and postgraduate degrees. We have not directly evaluated the impact of education loans on participation in tertiary education. In this section, we estimate the impact of education loan availability on contemporaneous enrolment. For information on enrolment in tertiary education and other individual and household level characteristics, we use data from three rounds of the NSS employment & unemployment survey conducted in 2004, 2009 and 2011. We use across district and over time variation in the extent of exposure to education loans to identify the impact of education loans on enrolment decisions.

The number of education loan accounts in an individual's district of residence at the time when she takes tertiary education enrolment decisions is an indicator of the extent of availability of education loans. However, the number of education loan accounts in our data is a cumulative number, the total number of education loan accounts in any year also includes the accounts of individuals who have completed education and are currently repaying their loans. We are now looking at contemporaneous enrolment and comparing individuals with varying but non-zero exposure to the program. The years that we include in our analysis are not far apart and some overlap in terms of the accounts included in the calculation of total accounts for each of these years is likely. We thus try to create a finer measure of the number of education loan accounts pertaining to the higher education age cohort in a particular year. We take the number of education loan accounts created in the last four years as a measure of the extent of exposure of the higher education age cohort in any year to education loans. While using the overall cumulative number would mean considering too many old accounts, taking differences over too short a period will lead us to exclude the

accounts pertaining to older individuals in the current higher education age cohort. Taking four year differences helps us strike a balance between the two. Four years is the maximum duration of any standard undergraduate course in India. The number of accounts created in the last four years would include not only the loan accounts relevant to those individuals who are currently taking the enrolment decision but also those relevant to older individuals who are currently enroled in tertiary education but had taken the enrolment decision earlier. Our results are robust to plausible alternative definitions of this measure of exposure, as demonstrated under robustness checks in this section.

The extent to which an individual is exposed to the improved availability of education loans depends on her district of residence and the time when she was in the higher education age cohort. We use these district and time level variations in exposure to identify the effect of education loan availability on enrolment decisions. The regression specification is as follows:

$$Y_{idt} = \alpha + \eta_d + \mu_t + \beta L_{dt} + \gamma \mathbf{X}_{idt} + \theta \mathbf{Z}_{dt} + \epsilon_{idt} \tag{3}$$

Y, the outcome variable of interest, indicates the enrolment status of an individual in the higher education age group. Y_{idt} takes value 1 if individual i (belonging to the higher education age group, 17-25 years) living in district d at year t is enroled in tertiary education[18] and zero otherwise. L_{dt} is the number of education loan accounts created in district d from year $t-4$ to year t per hundred population of the district, our primary independent variable of interest. $\mathbf{X}_{idt}$ is a vector of individual and household level controls such as sex, caste, sector (rural/urban), landholding and so on. To control for district level time invariant unobserved factors that might be correlated with both enrolment decision and education loan availability, we take district fixed effects. In equation (3), η_d represents the district fixed effects. We also control for time varying macroeconomic factors by including time fixed effects. μ_t represents the time fixed effects. There still might be factors varying across both district and time that

[18]Here by enrolment in tertiary education we refer to enrolment for attainment of any degree higher than higher secondary (high school).

can be correlated with both enrolment and access to education loan. We control for some such possible confounding factors: night light intensity per square km area of the district (a commonly used proxy for district GDP), population of the district and the number of bank branches in the district. $\mathbf{Z}_{dt}$ is a vector of these district level time varying controls. The summary statistics for the dependent variable, the main independent variable and district level controls have been presented in Table 10.

5.2 Results

5.2.1 Main Regression

Table 11 presents the results of regressing enrolment decision on the measure of education loan availability in the district in different specifications. Column 1 presents the results of a regression with education loan availability, household and individual level covariates, time fixed effects and district fixed effects. Column 2 presents our final specification and introduces night light intensity, population and bank branches as additional controls. The coefficient on the education loan variable is consistently significant and positive, with the magnitude falling only slightly after inclusion of the district and time varying controls. In our final specification, the probability of enrolment goes up by 0.066 for each unit increase in the number of new loan accounts per hundred population. The probability of enrolment goes up by 6.87% for each standard deviation increase in the loan availability as measured by the number of loans created in the district of residence per hundred population in the last four years.

5.2.2 Heterogenous Effects in Enrolment

Coefficients of demographic controls from the main regression of enrolment on education loans indicate lower enrolment probabilities for females vis-á-vis males, SC/ST or OBC vis-á-vis upper castes and rural population vis-á-vis urban population. In what follows we divide the data into groups and try to identify heterogeneous effects of educational loans on

enrolment for the disadvantaged sections.

Table 12 demonstrates the impact of loan availability separately on enrolment probabilities of male and female individuals. The magnitude of the coefficient of interest appears to be greater for male than for female. However, this difference is not statistically significant.

As presented in Table 13, we have divided our sample according to different caste groups and run the final regression specification for each of the groups. The coefficient of education loan accounts per hundred population is positive and statistically significant for all caste groups. The coefficients are not significantly different.

We next run the final specification separately for the rural and urban samples. The results have been summarised in Table 14. The coefficient on the number of new education loan accounts per hundred population is positive and statistically significant for the rural sector and insignificant for the urban sector. The difference in coefficients between the rural and urban sectors is statistically significant. This indicates that the program has had significant positive effect only in rural areas in terms of improving probability of enrolment.

The association between education loans and enrolment is closely tied with the relationship between family income and enrolment. We would have ideally liked to look at differential impact on different income groups. However, there is no information available in our data on a household's income. In rural areas, landholding can serve as a proxy of the household's economic status. Table 15 presents the regression results with the rural sample divided into quartiles of land holding. We find statistically significant positive effects in the second and third quartiles only. These results are suggestive of an inverted U type effect of the program, with the program being beneficial to those in the middle of the wealth distribution. Figure 2 presents a graphical depiction of the results. The statistically insignificant coefficient for the lowest land quartile might be indicative of the poorest households' inability to put their kids through the lower levels of education and of their limited access to education loans. For

the richest quartile, the insignificant effects might be indicative of fund constraints not being instrumental in the enrolment decisions of individuals from rich households.

5.3 Robustness Checks

The independent variable of interest is measured at the district level. We control for district and time fixed effects and some specific district level time varying factors in our regression specifications. The accuracy of our estimates of the effect of the education loan program on enrolment is still sensitive to correlations of other district level time varying factors with education loans and participation in tertiary education. Miss-specification of the measure of loan exposure also might lead to misleading results. In what follows, we describe robustness checks that were performed in order to mitigate these concerns.

The number of educational institutes is a district and time varying factor that could be correlated with both the number of education loans in a district and individuals' enrolment decisions. Hence not controlling for this number can potentially bias our results. Using data from the All India Survey of Higher Education (2014), we control for the number of colleges and universities in each district in a particular year. The result of this regression is presented in the second column of table 16. The coefficient of the education loan variable is still positive and statistically significant.

Though we cannot control for all district and time varying variables that can possibly bias our results, including interactions of the state and time dummies in the regression specification takes care of all time varying confounding factors at the state level. Since education is a subject in the concurrent list[19], there might be state specific policies that impact educational attainment and the take up of education loans. The third column of table 16 presents the results of the regression including state-time fixed effects. The coefficient of interest is

[19]In the Constitution of India the legislative section is divided into three lists: Union List, State List and Concurrent List. The Concurrent list includes subjects on which both the state government and the union government have legislative power.

positive and statistically significant, as before.

We have defined the extent of exposure to education loans in year t as the number of education loan accounts created in district of residence from year $t-4$ to year t per hundred population of the district. Our results should not be overly sensitive to the choice of this interval. In Table 17, we report results of regressions that use 3-year and 5-year differences[20] in loan accounts in the definition of the main independent variable. The coefficient of interest remains positive and statistically significant across both specifications.

Since we are using four year differences in education loan accounts to define the measure for the extent of exposure to the education loan program, there is an overlap between the periods we consider for calculating the numbers for 2009 and 2011. As outlined in Table 18, our results are preserved even if we run the main regression on data only from 2004 and 2011.

Since priority sector lending regulations apply to public sector banks and we would expect these banks to be more active in lending for education, our results should hold if we include only public sector bank accounts in our definition of the education loan variable. Results presented in the third column of table 18 show that the coefficient of interest is still positive and statistically significant.

6 Conclusion

In this paper, we have estimated the effects of education loan availability on individuals' tertiary education outcomes. We use evaluate the impact of loan availability on years of schooling by making use of variation in exposure to the education loan program by age cohorts and across districts. We find statistically significant positive effects of improved loan availability on years of schooling. For each unit increase in education loan accounts per

[20]The scheme was introduced in 2001. Hence the only measure available for 2004 is a three year difference, which is being used for all specifications.

hundred population in a district, the number of years of schooling achieved by the treatment cohort increased by 3.75 years. This signifies a 6.17% increase in years of education for each standard deviation increase in the number of education loan accounts per hundred population.

We use district and time level variation in exposure to the education loan program across district and over time to identify the effect of loan availability on enrolment in tertiary education. Our analysis yields statistically significant positive effects of years of schooling. For each unit increase in the number of loan accounts per hundred population, the probability of enrolment went up by 0.066. This translates into 6.87% improvement per standard deviation increase in the number of education loan accounts per hundred population.

In general, the positive effects of education loan availability are more pronounced for the relatively disadvantaged groups. The effects are stronger for female vis-á-vis male, SC/ST or OBC vis-á-vis the upper castes and rural population vis-á-vis urban population. However, the enrolment rates and years of schooling of these sections continue to be disproportionately lower. Our results indicate that any policy that focuses on increasing the accessibility of education loans for these sections will be a significant step towards removing existing divides and making tertiary education more inclusive.

In future research, we intend to develop a theoretical framework that explains the mechanism through which education loans impact tertiary education, keeping in mind the higher effects of loan availability on years of schooling for disadvantaged groups. The labour market implications of the positive effect of education loans on higher education are worth exploring. With granular data on education, the impact of loan availability on other aspects of tertiary education, like course choice, can be also be analysed.

References

Attanasio, O. and K. M. Kaufmann, Educational Choices, Subjective Expectations, and Credit Constraints, NBER Working Papers no. 15087, 2009.

Barr, A., "Enlist or enroll: Credit constraints, college aid, and the military enlistment margin", *Economics of Education Review* , 51 (2016) 6178.

Belley, Philippe, and Lance Lochner, "The Changing Role of Family Income and Ability in Determining Educational Achievement, *Journal of Human Capital*, 1 (2007), 3789.

Brown, Meta, John K. Scholz and Ananth Seshadri, A New Test of Borrowing Constraints for Education, *Review of Economic Studies*, 79 (2012), 511538.

Cameron, Stephen, and Christopher Taber, "Estimation of Educational Borrowing Constraints Using Returns to Schooling, *Journal of Political Economy*, 112 (2004), 132-182.

Cameron, Stephen, and James J. Heckman (1998): "Life Cycle Schooling and Dynamic Selection Bias: Models and Evidence for Five Cohorts of American Males," *Journal of Political Economy*, 106 (2), 262-333.

Cameron, Stephen, and James J. Heckman, The Dynamics of Educational Attainment for Black, Hispanic, and White Males, *Journal of Political Economy*, 109 (2001), 45599.

Carneiro, Pedro, and James J. Heckman, The Evidence on Credit Constraints in Postsecondary Schooling, *Economic Journal*, 112 (2002), 705-34.

de Ferranti, D., W. F. Maloney, G. E. Perry, I. Gill, J. Luis Guasch, Carolina Sanchez-Paramo, and Norbert Schady. 2003. "Closing the Gap in Education and Technology". *Latin American and Caribbean Studies*. Washington, DC:World Bank.

Duflo, Esther. 2001. "Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment", *American Economic Review*, 91(4): 795-813.

Dunlop, E., "What do Stafford loans actually buy you? The effect of stafford loan access on community college students", University of Virginia, 2012.

Dynarski, Susan M., "Does Aid Matter? Measuring the Effect of Student Aid on College Attendance and Completion, *American Economic Review*, 93 (2003), 279-288.

Galor, O. and J. Zeira, "Income Distribution and Macroeconomics," *Review of Economic Studies*, 60 (1993), 35-52.

Gurgand, Marc, Adrien Lorenceau, and Thomas Mélonio, "Student Loans: Liquidity Constraint and Higher Education in South Africa, Paris School of Economics working paper No. 2011-20, 2011.

Hanushek, E. A., and D. D. Kimko., "Schooling, Labor-Force Quality, and the Growth of Nations", *The American Economic Review* 90(5) (2000), 1184-208.

Hanushek, E. A., and L.Woessmann, "The Role of Education Quality for Economic Growth", Policy Research Working Paper No. 4122 (2007), World Bank,Washington, DC.

Henderson, J. Vernon, Adam Storeygard and David N. Weil (2012). Measuring Economic Growth from Outer Space. *American Economic Review*, Vol. 102(2), pp. 994-1028.

Krueger, Alan B., and M. Lindahl, "Education for Growth: Why and for Whom?" *Journal of Economic Literature* 39(4) (2001), 1101-36.

Kingdon, G., and M. Soderbom. 2007a. "Education, Skills, and Labor Market Outcomes: Evidence from Ghana." *Background paper prepared for the World Bank study, "Linking Education Policy to Labor Market Outcomes"*. Washington, DC.

Kingdon, G., and M. Soderbom. 2007b. "Education, Skills, and Labor Market Outcomes: Evidence from Pakistan." *Background paper prepared for the World Bank study,"Linking Education Policy to Labor Market Outcomes."* Washington, DC.

Kingdon, G. and J. Unni. (2001). "Education and Womens Labour Market Outcomes in India". *Education Economics* 9(2), 173-195.
Krueger, Alan B., and M. Lindahl, "Education for Growth: Why and for Whom?" *Journal of Economic Literature* 39(4) (2001), 1101-36.

Lange, Fabian, and R. Topel (2006), "The Social Value of Education and Human Capital", chapter 8 in: Eric Hanushek and Finnis Welch, eds, *Handbook of Economics of Education*, Vol.1, Elsevier, Amsterdam.

Lochner, Lance J., and Alexander Monge-Naranjo, The nature of credit constraints and human capital, *American Economic Review* 101 (2011), 24872529.

Manacorda, M., C. Sanchez-Paramo, and N. Schady, 2005. "Changes in Returns to Education in Latin America: The Role of Demand and Supply of Skills." *CEP Discussion Paper*

No. 712, Center for Economic Policy, London School of Economics and Political Science, United Kingdom.

Michalopoulos, S. and E. Papaioannou (2013) Pre-colonial Ethnic Institutions and Contemporary African Development, *Econometrica*, Vol. 81(1), pp. 113-52.

Montenegro, C., and H. Patrinos (2013), "Returns to Schooling Around the World". Background Paper for the World Bank World Development Report 2013.

Piketty, T. (1997). The Dynamics of the Wealth Distribution and Interest Rates with Credit Rationing, *Review of Economic Studies* 64, 173-189.

Pinkovskiy, M. L. (2013). Economic Discontinuities at Borders: Evidence from Satellite Data on Lights at Night Mimeo, Federal Reserve Bank of New York.
Rothstein, J. , and Rouse, C. E., "Constrained after college: student loans and early-career occupation choices",*Journal of Public Economics*, 95 (2010) , 149163.

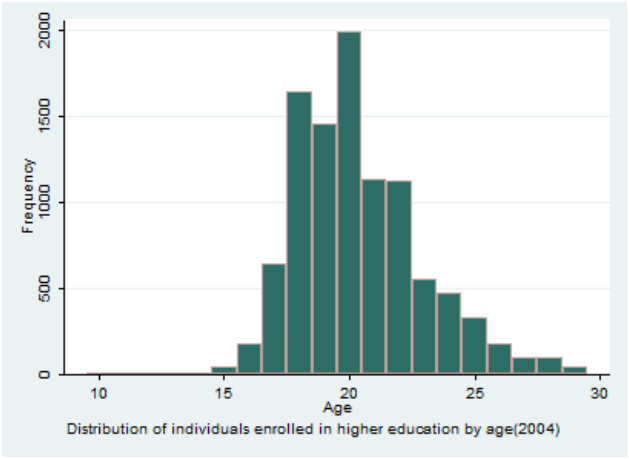
Schultz, T. P. 2003. "Evidence of Returns to Schooling in Africa from Household Surveys: Monitoring and Restructuring the Market for Education." *Economic Growth Center Discussion Paper No. 875*, Yale University, New Haven, CT.

Sianesi, B., and Van Reenen, J. "The returns to education: a review of the macro-economic literature." London: CEE Centre for the Economics of Education (LSE), 2000 (Discussion paper, DP 6).

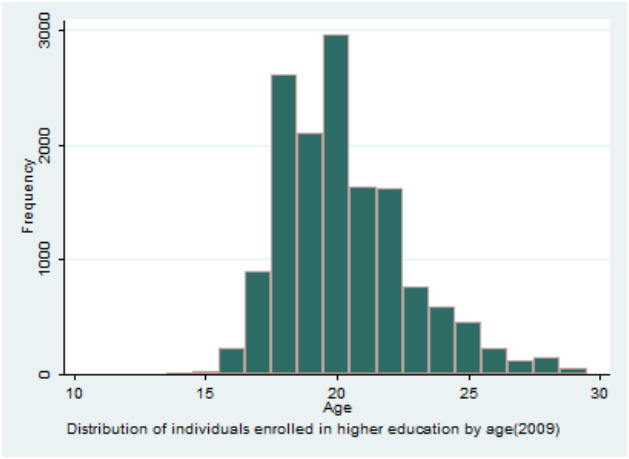
Solis, A., "Credit access and college enrollment", Berkeley, CA: Department of Agricultural and Resources Economics, University of California Unpublished, 2012.

Wiederspan, M., "Denying loan access: The student-level consequences when community colleges opt out of the Stafford loan program", *Economics of Education Review*, 51 (2016), 7996.

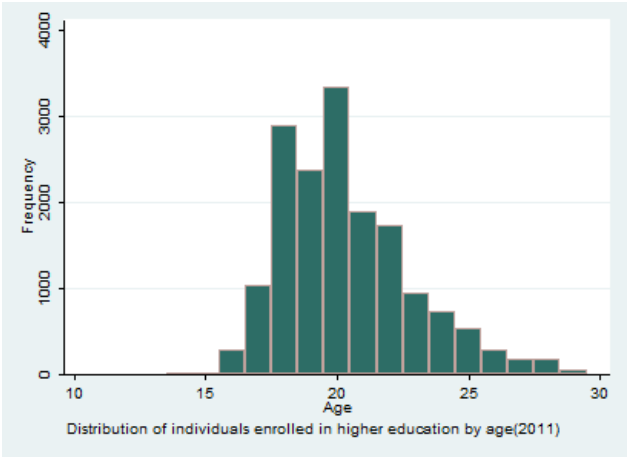
Figure 1: Enrolment in tertiary education by Age



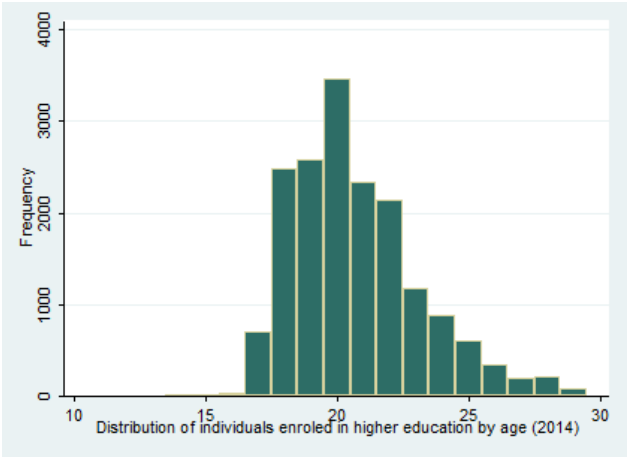
(a) 2004



(b) 2009



(c) 2011



(d) 2014

Figure 2: Effect of Education Loans on Enrolment: Heterogeneity Across Land Quartiles

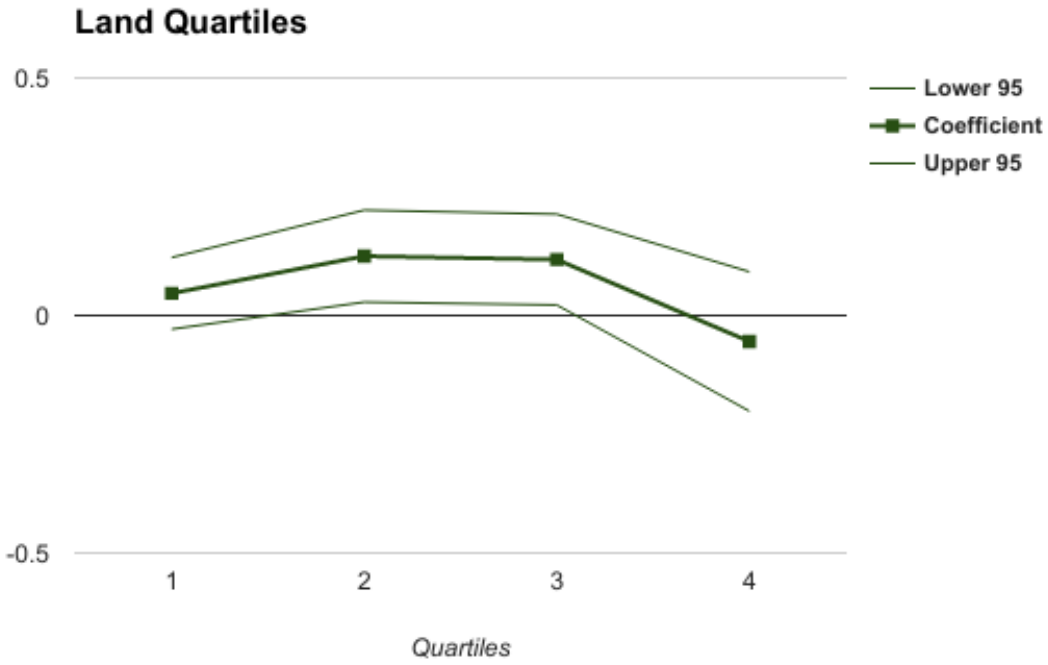


Table 1: Descriptive Statistics: Years of Schooling

	Mean	Std. Dev.
Education Loan Accounts (per Hundred Population in 2005)	0.048	0.066
Years of Schooling: 26-34 year olds (Measured in 2014)	7	5.41
Years of Schooling: 38-46 year olds (Measured in 2014)	5.144	5.27

Table 2: Effect of Education Loan on Years of Schooling

	(1) Years of Schooling
Program Intensity	3.753***
	(0.803)
Female	-2.232***
	(0.0715)
Urban	2.680***
	(0.106)
Scheduled Tribes & Scheduled Castes	-3.778***
	(0.1265)
Other Backward Castes	-2.146***
	(0.1266)
Demographic Controls	✓
Cohort Fixed Effect	✓
District Fixed Effects	✓
Observations	64362
R^2	0.323

Notes: In this table we present the estimate for the effect of loan availability on years of schooling achieved by an individual. Program intensity refers to the interaction between a dummy variable indicative of treatment status and the number of education loan accounts per hundred population in district of residence in 2005. The treatment dummy takes value one for the treatment cohort, 26-34 year olds in 2014, and zero for the control cohort, 38-46 year olds in 2014. Demographic controls include religion dummies. Standard errors are in parentheses, corrected for clustering at the district level. Cohort and district fixed effects have been controlled for.
$*p < 0.1, **p < 0.05, ***p < 0.01$

Table 3: Heterogeneous Effects in Years of Schooling: Gender

	(1) Male	(2) Female	(3) Overall
Program Intensity	1.632	5.597***	3.753***
	(1.289)	(0.965)	(0.803)
Female			-2.232***
			(0.0715)
Demographic Controls	✓	✓	✓
Cohort Fixed Effect	✓	✓	✓
District Fixed Effects	✓	✓	✓
Observations	30602	33760	64362
R^2	0.260	0.360	0.323

Notes: The sample has been divided according to gender. The first two columns present the effect of education loans on males and females respectively. Column 3 presents the overall results. Program intensity refers to the interaction between a dummy variable indicative of treatment status and the number of education loan accounts per hundred population in district of residence in 2005. The treatment dummy takes value one for the treatment cohort, 26-34 year olds in 2014, and zero for the control cohort, 38-46 year olds in 2014. Demographic controls include religion dummies, location (rural/urban) and caste. Standard errors are in parentheses, corrected for clustering at the district level.

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

Table 4: Heterogeneous Effects in Years of Schooling: Caste Groups

	(1) SC/ST	(2) OBC	(3) Others	(4) Overall
Program Intensity	7.548***	2.814**	-0.128	3.753***
	(1.234)	(1.354)	(1.742)	(0.803)
Scheduled Tribes & Scheduled Castes				-3.778***
				(0.1265)
Other Backward Castes				-2.146***
				(0.1266)
Demographic Controls	✓	✓	✓	✓
Cohort Fixed Effects	✓	✓	✓	✓
District Fixed Effects	✓	✓	✓	✓
Observations	16601	27082	20679	64362
R^2	0.297	0.304	0.336	0.323

Notes: Sample has been divided according to caste groups. The first column presents the results for Scheduled Castes and Scheduled Tribes. The second column presents the results for Other Backward Castes. Results for other upper castes are presented in column 3. The fourth column documents the overall results. Program intensity refers to the interaction between a dummy indicating treatment status and the number of education loan accounts per hundred population in district of residence in 2005. The treatment dummy takes value one for the treatment cohort, 26-34 year olds in 2014, and zero for the control cohort, 38-46 year olds in 2014. Demographic controls include religion dummies, gender and location (rural/urban). Standard errors are in parentheses, corrected for clustering at the district level.
$*p < 0.1, **p < 0.05, ***p < 0.01$

Table 5: Heterogeneous Effects in Years of Schooling: Sector (Rural/Urban)

	(1) Rural	(2) Urban	(3) Overall
Program Intensity	4.728***	3.004***	3.753***
	(1.328)	(0.693)	(0.803)
Urban			2.680***
			(0.106)
Demographic Controls	✓	✓	✓
Cohort Fixed Effect	✓	✓	✓
District Fixed Effects	✓	✓	✓
Observations	36642	27720	64362
R^2	0.292	0.245	0.323

Notes: Sample has been divided according to sector or location. The first and the second columns considers the rural and urban subsamples respectively. The third column presents the overall results. Program intensity refers to the interaction between a dummy variable indicative of treatment status and the number of education loan accounts per hundred population in district of residence in 2005. The treatment dummy takes value one for the treatment cohort, 26-34 year olds in 2014, and zero for the control cohort, 38-46 year olds in 2014. Demographic controls include religion dummies, gender and caste. Standard errors are in parentheses, corrected for clustering at the district level.
$*p < 0.1, **p < 0.05, ***p < 0.01$

Table 6: Effect of Education Loan on Years of Schooling: Falsification Test

	(1) Years of Schooling
Program Intensity	0.315
	(0.941)
Female	-2.635***
	(0.0677)
Urban	2.874***
	(0.106)
Scheduled Tribes & Scheduled Castes	-3.908***
	(0.1267)
Other Backward Castes	-2.436***
	(0.1312)
Demographic Controls	✓
Cohort Fixed Effect	✓
District Fixed Effects	✓
Observations	53226
R^2	0.3339

Notes: Program intensity refers to the interaction between a dummy variable indicative of treatment status and the number of education loan accounts per hundred population in district of residence in 2005. The treatment dummy takes value one for the pseudo-treatment cohort, 38-46 year olds in 2014 year olds in 2014, and zero the control cohort, 50-58 year olds in 2014. The results of this regression provide evidence against the results of the main regression being driven by pre-existing trends. Demographic controls include religion dummies. Standard errors are in parentheses, corrected for clustering at the district level.
$*p < 0.1, **p < 0.05, ***p < 0.01$

Table 7: Years of Schooling: Robustness Checks: Controlling for Higher Education Institutions and Bank Branches

	(1) Years of Education	(2) Years of Education	(3) Years of Education	(4) Years of Education
Program Intensity	3.753***	4.116***	2.389***	1.920**
	(0.803)	(0.921)	(0.752)	(0.807)
Baseline Institutions & Bank Branches		✓		
Institutions & Bank Branches at Age of Entry			✓	
Average Number of Institutions & Bank Branches				✓
Demographic Controls	✓	✓	✓	✓
District Fixed Effects	✓	✓	✓	✓
Observations	64362	64362	64362	64362
R^2	0.323	0.323	0.324	0.324

Notes: Program intensity refers to the interaction between a dummy variable indicative of treatment status and the number of education loan accounts per hundred population in district of residence in 2005. The treatment dummy takes value one for the treatment cohort, 26-34 year olds in 2014, and zero for the control cohort, 38-46 year olds in 2014. Demographic controls include religion dummies, gender, location (rural/urban) and caste. Standard errors are in parentheses, corrected for clustering at the district level. In this table we control for two possible confounders: bank branches and educational institutions. The first column controls presents the base specification. The second column controls for the interaction of cohort fixed effect and the pre-treatment baseline number of higher education institutions and bank branches in an individual's district of residence. The third column controls for the number of education institution and bank branches in an individual's district of residence at the time of her entry into the higher education age cohort. In the fourth column the average number of branches and institutions while an individual is in the higher education age cohort is controlled for.
$*p < 0.1, **p < 0.05, ***p < 0.01$

Table 8: Years of Schooling: Robustness Checks: Including Age and Household Fixed Effects

	(1) Years of Education	(2) Years of Education	(3) Years of Education	(4) Years of Education
Program Intensity	3.753***	4.471***	3.882***	5.066***
	(0.803)	(0.956)	(0.848)	(0.900)
Demographic Controls	✓		✓	
District Fixed Effects	✓		✓	
Cohort Fixed Effect	✓	✓		
Age Fixed Effect			✓	✓
Household Fixed Effect		✓		✓
Observations	64362	42333	64362	42333
R^2	0.323	0.841	0.330	0.845

Notes: Program intensity refers to the interaction between a dummy variable indicative of treatment status and the number of education loan accounts per hundred population in district of residence in 2005. The treatment dummy takes value one for the treatment cohort, 26-34 year olds in 2014, and zero for the control cohort, 38-46 year olds in 2014. Demographic controls include religion dummies, gender, location (rural/urban) and caste. Standard errors are in parentheses, corrected for clustering at the district level. The baseline specification with cohort and district fixed effects is presented in column 1. The second column controls for cohort and household fixed effects. The third column controls for district and age fixed effects. The specification in the fourth column includes age and household fixed effects.
$*p < 0.1, **p < 0.05, ***p < 0.01$

Table 9: Effect of Education Loans on the Probability of Attainment of Different Education Levels

	(1) Primary	(2) Secondary	(3) Higher Secondary	(4) Graduate	(5) Post Graduate
Program Intensity	-0.179*	-0.00266	0.0502	0.185***	0.0675***
	(0.0951)	(0.0578)	(0.0741)	(0.0520)	(0.0248)
Demographic Controls	✓	✓	✓	✓	✓
Cohort Fixed Effects	✓	✓	✓	✓	✓
District Fixed Effects	✓	✓	✓	✓	✓
Observations	64362	64362	64362	64362	64362
R^2	0.0441	0.0544	0.0488	0.0938	0.0633

Notes: This table presents results of a linear probability model estimated at various levels of education. For any level of education m, the dependent variable takes value one if the highest level of education attained by an individual is m. Program intensity refers to the interaction between a dummy variable indicative of treatment status and the number of education loan accounts per hundred population in district of residence in 2005. The treatment dummy takes value one for the treatment cohort, 26-34 year olds in 2014, and zero for the control cohort, 38-46 year olds in 2014. Demographic controls include religion, gender, location (rural/urban) and caste. Standard errors are in parentheses, corrected for clustering at the district level.
$*p < 0.1, **p < 0.05, ***p < 0.01$

Table 10: Descriptive Statistics: Enrolment

	2004		2009		2011	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Enrolment in Higher Education	0.09	0.281	0.15	0.359	0.18	0.3826
Number of Education Loan Accounts (in Last Four Years)	527.2888	1019.55	1805.275	3430.426	2389.384	4613.507
Number of Education Loan Accounts (per Hundred Population)	0.025	0.04	0.095	0.155	0.126	0.212
Night Light Intensity per sq. km	4.719	6.255	5.01	6.795	6.96	9.609
Number of Bank Branches	73.821	94.551	103.982	131.642	122.065	152.157
Number of Higher Education Institutions	35.839	46.519	60.406	74.673	69.462	83.721

Table 11: Effect of Education Loan on Higher Education Enrolment

	(1) Enrolment	(2) Enrolment
Education Loan Accounts (per hundred population)	0.0705*** (0.0170)	0.0659*** (0.0164)
Female	-0.0521*** (0.00298)	-0.0521*** (0.00298)
Scheduled Tribes & Scheduled Castes	-0.0444*** (0.00436)	-0.0446*** (0.00438)
Other Backward Castes	-0.0367*** (0.00417)	-0.0338*** (0.00418)
Urban	0.0378*** (0.00407)	0.0378*** (0.00406)
Demographic Controls	✓	✓
District Fixed Effects	✓	✓
Time Fixed Effects	✓	✓
District Level Time Varying Controls (Night Lights, Population and Bank Branches)		✓
Observations	188766	188766
R^2	0.145	0.146

Notes: This table presents the effects of education loans on enrolment in tertiary education. The first column controls for demographic factors and district and time fixed effects. The second column additionally controls for time varying variables, namely night lights, bank branches and population. The specification in the second column is our base specification. Education loan accounts refers to the number of education loan accounts per hundred population created in the last four years in an individual's district of residence. Demographic controls include religion dummies, land possessed, average adult education of household, number of dependent members, number of members in the higher education age group and household size. Standard errors in parentheses, corrected for clustering across districts.

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

Table 12: Heterogeneous Effects in Enrolment: Gender

	(1) Male	(2) Female	(3) Overall
Education Loan Accounts (per hundred population)	0.0672**	0.0642***	0.0659***
	(0.0263)	(0.0173)	(0.0164)
Female			-0.0521***
			(0.00298)
Demographic Controls	✓	✓	✓
District Fixed Effects	✓	✓	✓
Time Fixed Effects	✓	✓	✓
District Level Time Varying Controls (Night Lights, Population and Bank Branches)	✓	✓	✓
Observations	95131	93635	188766
R^2	0.157	0.153	0.146

Notes: The sample has been divided according to gender. The first column presents the results for males and the second column presents the effects of education loans on enrolment in females. The third column presents the results of the base regression in the overall sample. Education loan accounts refers to the number of education loan accounts per hundred population created in the last four years in an individual's district of residence. Demographic controls include religion, caste, location (rural/urban), land possessed, average adult education of household, number of dependent members, number of members in the higher education age group and household size. Standard errors are in parentheses, corrected for clustered at the district level.
$*p < 0.1, **p < 0.05, ***p < 0.01$

Table 13: Heterogeneous Effects in Enrolment: Caste Groups

	(1) SC/ST	(2) OBC	(3) Others	(4) Overall
Education Loan Accounts (per hundred population)	0.103*** (0.0365)	0.0523** (0.0210)	0.0992** (0.0435)	0.0659*** (0.0164)
Scheduled Castes & Scheduled Tribes				-0.0446*** (0.00438)
Other Backward Castes				-0.0338*** (0.00418)
Demographic Controls	✓	✓	✓	✓
District Fixed Effects	✓	✓	✓	✓
Time Fixed Effects	✓	✓	✓	✓
District Level Time Varying Controls (Night Lights, Population and Bank Branches)	✓	✓	✓	✓
Observations	46937	79488	62341	188766
R^2	0.128	0.130	0.183	0.146

Notes: Sample has been divided according to caste. The first column presents the results for Scheduled Castes and Scheduled Tribes. The second column presents the results for Other Backward Castes. Results for other upper classes are presented in column 3. Column 4 outlines the overall results. Education loan accounts refers to the number of education loan accounts per hundred population created in the last four years in an individual's district of residence. Demographic controls include religion dummies, caste, location (rural/urban), gender, land possessed, average adult education of household, number of dependent members, number of members in the higher education age group and household size. Standard errors are in parentheses, corrected for clustered at the district level.
$*p < 0.1, **p < 0.05, ***p < 0.01$

Table 14: Heterogeneous Effects in Enrolment: Sector (Rural/Urban)

	(1) Rural	(2) Urban	(3) Overall
Education Loan Accounts (per hundred population)	0.0666** (0.0274)	-0.00117 (0.00375)	0.0659*** (0.0164)
Urban			0.0378*** (0.00406)
Demographic Controls	✓	✓	✓
District Fixed Effects	✓	✓	✓
Time Fixed Effects	✓	✓	✓
District Level Time Varying Controls (Night Lights, Population and Bank Branches)	✓	✓	✓
Observations	120345	68978	188766
R^2	0.114	0.163	0.146

Notes: Sample has been divided according to sector or location.Columns 1 and 2 present results for rural and urban areas respectively. Column 3 presents the overall results. Education loan accounts refers to the number of education loan accounts per hundred population created in the last four years in an individual's district of residence. Demographic controls include religion dummies, caste, gender, land possessed, average adult education of household, number of dependent members, number of members in the higher education age group and household size. Standard errors are in parentheses, corrected for clustering at the district level.
$*p < 0.1, **p < 0.05, ***p < 0.01$

Table 15: Heterogeneous Effects in Enrolment in the Rural Sample: Land Quartiles

	(1) Q1	(2) Q2	(3) Q3	(4) Q4
Education Loan Accounts (per hundred population)	0.0472 (0.0398)	0.123** (0.0513)	0.138*** (0.0480)	-0.0338 (0.0717)
Demographic Controls	✓	✓	✓	✓
District Fixed Effects	✓	✓	✓	✓
Time Fixed Effects	✓	✓	✓	✓
District Level Time Varying Controls (Night Lights, Population and Bank Branches)	✓	✓	✓	✓
Observations	23998	27678	29336	39333
R^2	0.157	0.156	0.153	0.139

Notes: Rural sample has been divided into quartiles (weighted) of land possessed. The first four columns present the results of the regression of enrolment on education loans for each of these quartiles, with column 1 presenting the results for the first or the poorest quartile. Education loan accounts refers to the number of education loan accounts per hundred population created in the last four years in an individual's district of residence. Demographic controls include religion dummies, caste, gender, land possessed, average adult education of household, number of dependent members, number of members in the higher education age group and household size. Standard errors are in parentheses, corrected for clustering at the district level.

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

Table 16: Robustness Checks for Effect of Education Loan on Enrolment: Additional Controls

	(1) Baseline	(2) Institutes	(3) State×Time
Education Loan Accounts (per hundred population)	0.0659*** (0.0164)	0.0602*** (0.0164)	0.0680** (0.0293)
Demographic Controls	✓	✓	✓
District Fixed Effects	✓	✓	✓
Time Fixed Effects	✓	✓	✓
District Level Time Varying Controls (Night Lights, Population and Bank Branches)	✓	✓	✓
Observations	188766	188766	188766
R^2	0.146	0.146	0.147

Notes: Column1 contains the baseline estimates . In column 2, we add the number of higher educational institutes in an individual's district of residence as an additional control. in column 3 we add state×time fixed effects, allowing for state specific trends. Education loan accounts refers to the number of education loan accounts per hundred population created in the last four years in an individual's district of residence. Demographic controls include religion dummies, caste, gender, location (rural/urban), land possessed, average adult education of household, number of dependent members, number of members in the higher education age group and household size. Standard errors are in parentheses, corrected for clustering at the district level.
$*p < 0.1, **p < 0.05, ***p < 0.01$

Table 17: Robustness Checks : Alternative Definitions of the Measure of Loan Availability

	(1) Baseline	(2) 3-year differences	(3) 5-year differences
Education Loan Accounts (per hundred population)	0.0659*** (0.0164)	0.0754*** (0.0291)	0.0577*** (0.0216)
Demographic Controls	✓	✓	✓
District Fixed Effects	✓	✓	✓
Time Fixed Effects	✓	✓	✓
District Level Time Varying Controls (Night Lights, Population and Bank Branches)	✓	✓	✓
Observations	188766	133686	187641
R^2	0.146	0.148	0.148

Notes: Column1 contains the baseline estimates where education loan accounts refers to the number of education loan accounts per hundred population created in the last four years in an individual's district of residence. In this table we demonstrate the robustness of our results to plausible alternative definitions for the independent variable. In column 2, we take 3-year differences in loan accounts and in column 3 we we take 5-year differences in loan accounts. Demographic controls include religion dummies, caste, gender, location (rural/urban), land possessed, average adult education of household, number of dependent members, number of members in the higher education age group and household size. Standard errors are in parentheses, corrected for clustering at the district level.
$*p < 0.1, **p < 0.05, ***p < 0.01$

Table 18: Robustness Checks : Excluding 2009 and Loan Accounts for Public Sector Banks

	(1) Baseline	(2) 2004 & 2011	(3) Public Sector Banks
Education Loan Accounts (per hundred population)	0.0659*** (0.0164)	0.0737*** (0.0185)	0.0551*** (0.0136)
Demographic Controls	✓	✓	✓
District Fixed Effects	✓	✓	✓
Time Fixed Effects	✓	✓	✓
District Level Time Varying Controls (Night Lights, Population and Bank Branches)	✓	✓	✓
Observations	188766	188766	188766
R^2	0.146	0.148	0.146

Notes: Column1 contains the baseline estimates.In column 2, we include data corresponding to 2004 & 2011, thus dropping 2009. In column 3 we include loan accounts from only public sector banks thus dropping private sector banks from our sample. Education loan accounts refers to the number of education loan accounts per hundred population created in the last four years in an individual's district of residence.Demographic controls include religion dummies, caste, gender, location (rural/urban), land possessed, average adult education of household, number of dependent members, number of members in the higher education age group and household size. Standard errors are in parentheses, corrected for clustering at the district level.
$*p < 0.1, **p < 0.05, ***p < 0.01$