Improving First-Generation College Students' Education and Labour Market Outcomes: Impact Evaluation of an Inclusive Policy in India

Gopinath Annadurai * Soham Sahoo *

December 7, 2023

Abstract

First-generation graduates (FGGs), the first in their families to attend college and representing a growing segment in higher education, encounter unique aspirations and challenges. Despite their significance, there has been limited research and policy interventions on FGGs. This study evaluates the First-Generation Graduate Scholarship (FGGS) scheme, launched in 2010 in the Indian state of Tamil Nadu, which aims to waive tuition fees for FGGs pursuing technical education in engineering, medicine, and agriculture at the undergraduate level. Our analysis focuses on FGGs aged 17–22 and assesses the program's impact on professional course enrollment, stream choice, and subsequent labour market outcomes. We employ a Difference-in-Differences (DID) model using multiple rounds of the National Sample Survey. Our findings reveal a notable 3.6 percentage point increase in professional course enrollment in response to the policy; this effect translates to around 43% increase over the mean enrollment rate in professional courses. The treatment also significantly affects the beneficiaries' stream choice and graduate-degree completion rates in favour of professional courses. Various robustness checks, including synthetic-DID analysis, event studies, and placebo tests, affirm the program's effectiveness in promoting enrollment in professional courses. Beyond academic outcomes, our study explores labour market consequences, demonstrating that the policy led to a shift towards service-sector employment among FGGS beneficiaries, accompanied by reduced engagement in agriculture-related work, decreased casual employment, and an increased propensity for active job-seeking.

Keywords: First-generation graduates, Professional enrollment, Tuition fee waiver
JEL Codes: I23, I24, I28, J24

*Indian Institute of Management Bangalore. Correspondence to: gopinath.annadurai21@iimb.ac.in

1 Introduction

First-generation graduates (FGG) are the first in their families to attend college. This growing segment within higher education has distinctive aspirations, incentives, and challenges (Hsiao, 1992). Typically hailing from working-class backgrounds or belonging to ethnic minorities or women, these newcomers to higher education often grapple with unique hurdles on their path to obtaining a degree. These hurdles include conflicting responsibilities, unfounded expectations, and a lack of preparedness and support (Padrón, 1992).

While students with college-educated parents view their enrollment in higher education as a natural and anticipated step towards personal and professional success, it is a different story for others. For this second group, enrolling in higher education is a conscious effort to enhance their social, economic, and career status. Among these students, many are breaking new ground in their families by pursuing education beyond high school, marking them as "first-generation students". Higher education presents opportunities and risks for these individuals, as it signifies a departure from long-standing family traditions (Nuñez et al., 1998).

Higher education policies must be adapted to align with the requirements of first-generation graduates. Nevertheless, there has been limited investigation into this issue in the Indian context with regard to both research efforts and policy initiatives. A unique development in this sphere is the initiative embarked upon by the state of Tamil Nadu. This initiative recognizes first-generation students as a distinct cohort and has introduced a scholarship program expressly tailored to address the needs of this demographic.

This paper evaluates the effect of the first-generation graduate scholarship (FGGS) scheme implemented in Tamil Nadu since 2010. The scheme aims to waive "tuition fees" for first-generation graduate students taking technical education courses such as engineering,

medicine, and agriculture at the undergraduate level. The sole criterion for eligibility is that the student should be the first graduate of the family, irrespective of gender, caste, religion, and income status of the family. The data shows that 1,91,268 students have benefitted with Rs. 402.69 crore spent on this program in the academic year 2018-2019 alone. FGGS scheme is one of the most extensive affirmative action programs on higher education in Tamil Nadu, and around 10% of the total budget for the state's higher education is spent on this program every year since the inception of the scheme.

We examine the program's effects on professional course enrollment and stream choice among first-generation college students in this context. Further, we also explore the labour market consequences of the program on the target group. We use the district representative pooled household data from the NSS rounds 61, 64, 66, 68, 71, and 75 to estimate the impact on enrollment. Professional enrollment is considered in terms of graduation current enrolment in the engineering, agriculture and medical courses, and stream choice is measured as professional enrollment conditional on being enrolled in any undergraduate courses.

We use the exogenous timing in the program's implementation for identification and use difference-in-differences (DID) regression to assess the intent-to-treat (ITT) program effect. We have selected the neighbouring states of Kerala and Karnataka as the control states, considering the socio-economic and cultural similarities with the treatment state, Tamil Nadu. Here, we compare the trends in professional enrollment and stream choice of first-generation college-goers in the age cohort of 17–22 years in the treated state with that of the first-generation college-goers in the same cohort from the control states. We consider years 2004, 2007 and 2009 as the pre-intervention period and years 2011, 2014 and 2017 as the post-intervention period in the pooled data. We also conduct a battery of robustness tests, including synthetic-DID, event study, and placebo analysis. We test for the parallel trends assumption in all of our DID models.

Our findings suggest that the FGGS scheme increased professional graduate enrollment among first-generation individuals, and the scholarship made more people from the target group shift to professional courses from non-professional courses. The results are similar even after controlling for individual and household-level factors and unobserved heterogeneity at the state and district levels. Our findings reveal a notable 3.6 percentage point increase in professional course enrollment in response to the policy; this effect translates to around 43% increase over the mean enrollment rate in professional courses. The treatment also significantly affects the beneficiaries' stream choice decision by increasing the likelihood of choosing a professional course instead of a non-professional course by 8 percentage points (i.e., 32% over the mean).

We also estimate the labour market consequences of the FGGS scheme. The pooled data for this analysis was sourced from NSS rounds 66, 68, and PLFS (Periodic Labour Force Survey) 17-18 and 18-19. First, we find that the scheme had a positive and significant impact in terms of increasing the likelihood of having a professional graduate degree among the beneficiaries. This implies that the scheme was not only successful in enrolling students in professional degree courses but also in improving the completion rates among the targeted group. Then, we observe that the scholarship program steers first-generation graduates towards service-sector employment while discouraging engagement in agriculture and related sectors. We also observe a decrease in casual employment and an increase in the propensity for active job-seeking.

We present suggestive evidence indicating a decrease in educational expenditure and the proportion of educational spending among First-Generation Graduates (FGGs) who enrolled in professional courses in Tamil Nadu subsequent to the implementation of the FGGS scheme. Furthermore, our analysis explores the consequences of professional graduation on earnings

and social welfare, revealing suggestive evidence of an augmentation in monthly per capita expenditure among FGGs.

We perform a series of robustness checks to examine the sensitivity of our estimates. We use the synthetic difference-in-difference (S-DiD) with more donor states, along with Kerala and Karnataka, comparable to the treatment state, Tamil Nadu. We have used 18 states in the pre-intervention period to form our donor pool and create a synthetic state that looks similar to Tamil Nadu. The S-DiD estimates are similar to our main results. We also employed an event study and placebo test for pre-intervention periods. The results show no significant pre-intervention trends, implying that the parallel trends assumption is likely to be satisfied in our context. Our estimates suggest a robust and significant impact of the program on enrollment throughout these sensitivity analyses. The FGGS scheme increased the professional graduate enrollment among first-generation graduates, and the scholarship made more people from the target group shift to professional courses from non-professional courses.

Our heterogeneity analysis indicates that the program has a more pronounced effect on the upper strata within the cohort of First-Generation Individuals. Specifically, the program demonstrates a greater impact on First-Generation Graduates (FGGs) who are male, belong to the Other Backward Classes (OBC), or reside in urban areas, as compared to their respective counterparts.

The rest of the paper is organized as follows. In Section 2, we describe details about the program. Section 3 describes the dataset's main variables and presents some descriptive analysis. Section 4 explains the empirical model. The results of the analysis are discussed in Section 5. Section 6 concludes the paper.

2 Background

2.1 First-generation graduates

The FGGS scheme identifies and targets the state's first-generation graduates (FGG). According to this program, the scholarship is eligible for the student, given that none of his/her household members are graduates. Students who come from no-graduate household members family are first graduates. Though treating first-generation graduate individuals as a group/ category for affirmative action is new for Indian states, the term is widely used in Western academia. Extant literature on first-generation graduates in education focuses on various topics such as college and stream choices, educational outcomes, comparison with other category students, cultural conflict and mental health.

First-generation graduation is a growing and influential category in American educational institutes (Pascarella et al., 2004). First-generation students differ from their traditional peers in entering characteristics and college experiences (Terenzini et al., 1996). Moreover, they are likely to enter with lesser preparations than the traditional groups (Thayer, 2000). They do not know how to finance a college education to complete basic admissions procedures (Vargas, 2004). Furthermore, they do not know how to connect career goals and educational requirements. They face some unique challenges before and after the admission process (Striplin, 1999).

In the realm of academia, it has been observed that a considerable proportion of first-generation students encounter apprehensions within the educational framework, often grappling with a lack of clarity regarding its adaptability and rigidity (Padrón, 1992). Furthermore, these students exhibit a higher attrition rate when compared to their non-first-generation counterparts (Nuñez et al., 1998).

Hence, higher education policies must be tailored to meet the needs of first-generation graduates. However, these matters have not received extensive exploration within the Indian context, both in terms of research endeavours and policy analyses. An exciting development in this context is the pioneering initiative undertaken by the state of Tamil Nadu, which has recognized first-generation students as a special/separate group and introduced a scholarship program specifically designed to address the needs of this demographic.

2.2 FGGS Scheme

The scheme has been in place in the state of Tamil Nadu from the academic year 2010-11 through the G.O.(Ms)—no: 85, Higher Education (J2) Department. The scheme aims to give the waiver of "tuition fees" for first-generation graduate students who took up technical education courses via the single window counselling system of the Government. According to the G.O., the sole criteria for selecting students is that the student should be the first graduate of the family with no regard to the caste and income status of the student and his/her family for eligibility. The scheme has been extended to Medical, Dental, Engineering, Veterinary, and Agricultural colleges.

The nature of the institution could be Government, Govt- Aided, and private self-financing institutions. Though admission for Law colleges has not been made through a single-window system, later, as a special consideration, the scheme has also been extended to law college first graduates in successive years. In addition, those first graduate diploma students who take up technical graduation courses such as BE, B.Tech, etc., through lateral entry are also considered eligible in this scheme. Upon the verification and submission by the institution to the government departments, the Government will directly pay the tuition fees to the respective institutions. The State Government has fixed the amount of tuition fees in Government colleges. The amount for private self-financing institutions has been fixed by a

committee constituted by the Government under a retired judge. Accordingly, the tuition fee waiver amount would be fixed based on the committee's recommendation.

The importance given by the State government for this scheme in the Higher education department is evident in the budgetary allocations of the state. Out of Rs. 5,052 crores allotted for the Higher Education Department in the Budget Estimates 2020-2021, a sum of Rs.506.04 crore has been provided for this program. For 2017-18, a sum of Rs.434.56 crore has been sanctioned, and 2,17,396 students have benefitted. For 2018-19, data shows that 1,91,268 students have benefitted, with Rs.402.69 crore being spent on this program.

3 Data

3.1 Data Description

To assess the impact of the First Generation Graduate Scholarship (FGGS) scheme on professional enrolment, we used pooled household data from the NSSO, a district-representative independent household survey conducted on various topics. The surveys are conducted regularly by the Indian premier data collecting agency, the National Sample Survey Organization (NSSO). The data is sourced from NSS rounds 61, 64, 66, 68, 71, and 75 for our study. The 61st round survey on 'employment and unemployment' taken in 2004, 64th round survey on 'Participation and Expenditure in Education' taken in 2007, 66th round survey on 'employment and unemployment' taken in 2009, 68th round survey on 'household consumer expenditure' taken in 2011, 71st round survey on 'Social Consumption: Education' taken in 2014, 75th round survey on 'Household Consumer Expenditure' taken in 2017 were used for the study.

Each survey covers 0.3 to 0.7 million individuals. For instance, A sample of 4,45,960

persons, from 63,318 rural households and 37,263 urban households spread over the country, was surveyed in the NSS 64th round (July 2007 - June 2008). Moreover, a sample of 3,10,827 individuals, from 36,479 rural households and 29,447 urban households spread over the country, was surveyed in the NSSO 71st round.

Though these surveys are conducted on various topics, all of them have common information like socioeconomic variables of the households, education details of every member of the household, and participation of individuals aged 5-29 years in the country's education system. The period in which the surveys took place is apt to do a comparative analysis of before and after the implementation of the FGGS scheme and compare the same with a set of other neighbouring states as well. The variable of interest for our study is the enrolment in professional graduate degrees and the stream choice of students enrolled at the graduation level.

3.1.1 Main variables

Since the policy is aimed at graduation degrees, we define the age of 17-22 as an undergraduate college-going cohort. And the individuals from households with no prior graduated members are eligible for the scholarship, the potential candidates for the FGGS scheme. Hence, we define the individual from the age group of 17 to 22 as a potential candidate if none of his household members in the above 23 age group are graduates in any field. We limited our analysis only to these potential candidates, around 87% in the 17-22 age cohort.

Our main variable of interest is whether the current enrollment of a FGI is in a professional undergraduate program. This study defines the Engineering, Medicine, and Agriculture degree as a professional course. Since this scholarship is only for a graduate degree, we did not consider the certificate, diploma, or post-graduate courses as graduation here.

NSSO records the household's current enrollment status of persons aged 5-29. It also records the education particulars on the basic course of those aged 5-29 years who attend primary level and above. The course code in the education particulars indicates the type of the courses, such as humanities, science, commerce, medicine, engineering, agriculture, management, etc. We have created a binary outcome variable, which indicates the enrollment in the professional under graduation degree, using the current enrolment status and course code. The variable professional enrolment, "prof_enrolment", is taken as one if the individual of age cohort 17-22 is currently enrolled in engineering/medical/agriculture under graduation degree. And 0 otherwise.

3.1.2 Other variables

We have used a set of individual, household, village and district-level information as independent variables in the regressions, which are possible confounders of professional graduate enrolment. Household characteristics were controlled through several indicators: Highest education qualification of household members other than the potential candidates, religion and social category of the household, and total members within the household. Individual-level characteristics include the age, sex, relation of the person to the head of the household and the individual's marital status.

The sector of the household, whether the household is in a rural or urban area, is used as a village-level control. District-level factors include the total number of colleges in the district and the proportion of professional colleges among the colleges in the district. Controlling for these characteristics enabled us to get an unbiased estimate of the impact of the FGGS scheme.

We use the neighbouring states Kerala and Karnataka as a comparison state to Tamil Nadu to evaluate the program. The sample for our main analysis for professional enrolment is defined as 17-22-year individuals from first-generation households in Tamilnadu, Kerala, and Karnataka. The dataset encompasses a total of 37,261 potential candidates. The summary statistics for the enrolment sample are reported in the All sample column of Table 1.

Approximately 8.5% of individuals in the sample are enrolled in professional courses, with a standard deviation of 0.278, suggesting some degree of variability in enrollment rates. The data includes individuals from different caste categories, with the majority belonging to the "Other Backward Class" (62.9%) and "Hindu" religion (76.5%). The dataset is pretty evenly split between rural (57.4%) and urban (42.6%) residents, and 47.3% of the sample is female. National weights were used in our analysis.

3.2 Stream choice

Another variable of interest for our study is the stream choice of students enrolled at the graduate level. The FGGS scheme, which aims only at professional courses, may impact the stream choice of students who enrol in graduation degrees. The scheme incentivises students to choose professional degree courses over non-professional degree courses. Generally, professional courses like engineering, medicine and agriculture are more expensive than non-professional courses such as arts, humanities, commerce etc. Since the FGHHS scheme substantially reduces the professional course cost, more people can afford the professional degree now.

In order to estimate the impact of the FGGS scheme on stream choice at the graduation level, we construct a binary variable indicating enrollment in the professional graduate course versus enrollment in the non-professional graduate course. The outcome variable is defined only for the potential candidates (none of his household members above 23 age group

are graduates in any field) of the college-going age group 17-22 years and who enrolled in a graduate degree. If the individual is enrolled in a professional degree, the binary variable takes 1, and if enrollment is in a non-professional degree, the value takes 0.

Table 1-column B provides the summary statistics of the variables used in stream choice analysis. The final sample for stream choice analysis considers only the potential candidates of the scheme who are currently enrolled in undergraduate degrees and within the age group of 17 to 22 years. Thus, we consider 12,892 potential candidates who were enrolled in any undergrad degree courses for our analysis. 24.5% of individuals in this subset are pursuing professional courses.

We observed some differences in the conditional enrollment sample from all samples. The proportion of males in the sample increased from 52% to 74% in the conditional sample, indicating the high magnitude of gender difference in the higher studies. The proportion of SC ST slightly reduced, and OBC and Other castes (Upper caste) slightly increased. The rural proportion is also lower in the conditional sample. We observed differential behaviour among individuals from different social groups, genders, religions and sectors in the general higher studies enrolment.

3.3 Labour market outcomes

3.3.1 Data

Changes in enrollment and stream choice at the graduation level affect the labour market outcomes of the individual after graduation. Professional graduation can significantly influence an individual's career prospects and earning potential. The labour market returns to education, which researchers widely study. These returns are typically measured in terms of improved labour market outcomes, such as higher earnings and better job prospects. Edu-

cation can also enhance an individual's employability. We want to estimate the changes in the labour market in terms of the type of jobs and sector because of the FGGS scheme.

The data for this research study was sourced from NSS rounds 66, 68, and PLFS (periodic labour force survey) 17-18 & 18-19, conducted by the National Statistical Office (NSO) of India. The NSS 66th (July 2009 - June 2010) and 68th (July 2011 - June 2012) rounds of household surveys focus on employment and unemployment trends at both national and state levels. The National Statistical Office (NSO) introduced the Periodic Labour Force Survey (PLFS) in April 2017 to address the need for more frequent labour force data. The PLFS datasets were instrumental in providing more frequent estimates of key employment and unemployment indicators, such as the Worker Population Ratio, Labour Force Participation Rate, and Unemployment Rate. We utilised two PLFS datasets, PLFS 17-18 and PLFS 18-19, in our analysis. All of them are an all-India household survey and district representative. We have generated pooled household data from these datasets for our analysis.

We are interested in the period when the target beneficiaries enter the labour market. We have defined the college-going age group as 17-22 in the enrolment analysis. Professional courses take 4 to 7 years to complete, depending on the degree they are enrolled in. Considering the graduation duration, we define the age group of 25 to 29 as the potential period for the beneficent cohort to be in the labour market. Hence, we define the individual from the age group of 25 to 29 as a potential policy candidate if none of his household members in the above 29 age group have graduated in any field. Even here, we have limited our sample only to these potential candidates.

3.3.2 Variables

We have derived our outcome variables to measure the individual's employment status from the usual principal activity status of household members. We have recorded the principle activities into multiple binary variables indicating whether the individual is self-employed, regular salaried, casual, or seeking employment. These variables are defined only by the potential candidates of the age group 25- 29 years.

The variable self-employment takes the value 1 if the principal activity of the individual is worked in the household enterprise (self-employed) as an own account worker or employer or worked as an unpaid family worker; otherwise, it takes 0. The variable regular-employment takes value 1 if the principal activity of the individual is worked as a regular salaried/ wage employee; otherwise, it takes 0.

The variable casual employment takes value 1 if the principal activity of the individual is worked as casual wage labour either in public works or other types of work; otherwise, it takes 0. Likewise, the variable seeking employment takes the value 1 if the principal activity of the individual is "did not work but was seeking and (or) or available for work"; otherwise takes the value 0.

We clubbed every other response in "the principal activity status," such as attending educational institutions, attending domestic duties only, were also engaged in free collection of goods, sewing, tailoring, weaving, etc. for household use, rentiers, pensioners, remittance recipients, etc., not able to work due to disability, as the binary variable for "Not in labour force".

We are also interested in estimating the impact on the labour market in terms of variation in the industry occupation. The survey records the type of industry for those who are

employed. The National Industrial Classification (NIC) 2008 codes are used for the classification of the economic activity of the individual. Using this code, we have created three binary variables: Primary, Secondary and Service sector. These variables are defined only for the potential candidates of the age group 25- 29 years and also employed.

The National Classification of Occupations (NCO) determines the skill level associated with each occupation by considering factors such as academic and technical qualifications, experience prerequisites, and the job's inherent characteristics. This assessment includes whether the job entails administrative, managerial, and supervisory responsibilities or is of a subordinate or repetitive nature, considering the context in India. We've directly utilized the four skill levels outlined in the NCO 2004 and 2008 as a binary variable. This binary variable is employed to estimate the impact of FGGS concerning the occupation's skill level for individuals who are employed.

We also created the variable professional graduate, "prof_grad", to estimate the first-order effect in the new sample. The variable, "prof_grad", takes value 1 if the potential candidate of age cohort 25-29 is graduated with an engineering/medical/agriculture degree and 0 otherwise.

Table A.3 in the annexure summarises the variables used to estimate the labour market outcomes.

4 Empirical Strategy

4.1 Difference-in-Differences (DID)

The survey timeline allowed us to look at the conditions before and after implementing the FGGS and compare the same with a set of other neighbouring states. We looked at the status of professional graduation enrolment in Tamil Nadu before and after implementing the FGGS Scheme and compared the same with the neighbouring states of Kerala and Karnataka through the Difference in Differences (DiD) method.

We considered the potential candidates, aged 17 to 22 years, from the no graduate household in Tamil Nadu as a treatment group. And the same group from Kerala and Karnataka together as a control group, who would not be exposed to the FGGS. Our sample period began in 2005, 6 years before the program's implementation. Surveys 61, 64, and 66 were taken 6 and 4,2 years before implementing the program. Survey 68 was taken in 2011, and the 71st and 75th rounds were taken 3 and 6 years after implementing the program, respectively (2014 & and 2017). Hence, the data from 61, 64, and 66 rounds are used as a control cohort, and 68, 71 and 75 rounds are used as a treatment cohort.

We made use of the Difference-in-Difference (DID) regression methodology first to get the Intent-to-Treat (ITT) estimate of the programme in increasing professional enrolment among the 17-to-22 age group, and we estimated the changes in the stream choice, conditional on the enrolment in any graduate program using the same model. The first difference compared the outcome across the treatment group of Tamil Nadu before and post-implementation of the program. The second difference compares the same across the state.

The Difference in -Differences estimate of exposure to the FGGS program is given in Eq.(1). We estimate the equation separately for professional enrollment and stream choice.

$$Y_{idst} = \alpha_0 + \beta_1 Post_t + \beta_2 TN_s + \beta_3 (Post_t \times TN_s) + \beta_4 X_{idst} + \mu_{ds} + \gamma_t + \epsilon_{idst} \qquad (1)$$

Here Y_{idst} is the outcome variable of interest, indicates enrollment in the professional degree, for individual, i from district, d, state, s, and time, t. The 'post' dummy takes the value of 1 if the year is 2011, 2014 or 2016 and 0 if the year is before 2010. 'TN' is a dummy that takes the value 1 if the i-th observation comes from Tamil Nadu; otherwise, 0. β_3 is the causal estimate of the impact of FGGS on the outcome variables. The confounding individual, household, village and district characteristics that may affect the outcome variables are controlled by X_{idst}. We also controlled for district and year fixed effects (μ_{ds} and γ_t respectively) to control for the secular state and year-level changes that might affect the outcome variables. All the standard errors are clustered at the district level. Even though the policy was implemented at the state level, the number of clusters is too small if we cluster the standard errors at the state level. To overcome this difficulty, we have also used the wild cluster bootstrap method and the corresponding p values are reported within parentheses in the regression tables.

4.2 Pre-Trend analysis

Before estimating the treatment effect, we examined pre-intervention trends to assess whether any patterns before intervention might have systematically influenced the treatment or control state. Prior to implementing the FGGS scheme, there could have been divergent trends in education and employment between the treatment and comparison states. To assess the pre-trends assumption, we compare the trends in the outcome variables between the treatment and control groups before the implementation of the treatment.

The Difference in -Differences model for pre-trend analysis is given in Eq.(2).

$$Y_{idst} = \alpha_0 + \beta_1 time_t + \beta_2 TN_s + \beta_3 (time_t \times TN_s) + \beta_4 X_{idst} + \mu_{ds} + \gamma_t + \epsilon_{idst} \quad (2)$$

Here, the model is similar to equation 1. The sample is confined to the pre-treatment period, with $time_t$ assuming values 1, 2, and 3 for the years 2004, 2007, and 2009, respectively. The parameter estimate of β_3, if statistically insignificant, ensures non-rejection of pre-trends.

4.3 Event study

We use a Difference-in-Difference (DID) event study or a Dynamic DID model as a robustness check. The event study is a visual representation showing the regression's point estimates and confidence intervals for each period before and after the treatment period. In DiD event study, the estimated values represent the average differences between the treated and control groups. This strengthens the validity of assuming parallel trends.

The Difference in -Differences (DiD) event study estimate is given in Eq. (3).

$$Y_{idst} = \alpha_0 + \sum_{t \neq 2009} \beta_t (TN_s \times Year_t) + \lambda X_{idst} + \mu_{ds} + \gamma_t + \epsilon_{idst} \quad (3)$$

Here, Y_{idst} is the outcome variable of interest, indicates enrollment in the professional degree, for individual, i from district, d, state, s, and time, t. Control variables X_{idst}, district and year fixed effects (μ_{ds} and γ_t respectively) are the same as Eq.(1). And the standard errors are clustered at the district level. $Year_t$ takes the value of 1 if the observation i belongs to year t, where $t \in \{2004, 2007, 2011, 2014, 2017\}$, the reference group being the observation belonging to the year 2009.

4.4 Synthetic-DiD

As an additional robustness check, we estimate effects using the synthetic difference-in-differences method proposed by Arkhangelsky et al. (2021). In this method, the impact of the FGGS Scheme on the potential candidates (First graduate individuals) in entering a graduate program is analysed using counterfactual estimation. The situation of how first-graduate individuals would access professional graduate programs in the absence of the scheme in Tamil Nadu is analysed by constructing a synthetic Tamil Nadu state.

The S-DiD complements both the standard SCM and the difference-in-differences approaches. Standard synthetic control methods (SCM) construct a counterfactual as a weighted average of comparison units - states. The SCM consists of a weighted regression with time-fixed effects but no unit-fixed effects. The difference-in-differences can be thought of as an unweighted regression with both time and unit fixed effects. The S-DiD estimator integrates these approaches by calculating weights for periods, with weights calculated to achieve a balance between pre- and post-program periods (Abadie, 2021; Arkhangelsky et al., 2021).

This method fits well with our aims as it also allows studying the FGGS scheme's effects on potential Tamil Nadu candidates by constructing counterfactuals from the other states and the pre-intervention period. The method does not rely on a parallel-trends assumption, and it corrects for both unit and time weights, typically assigning larger weights to the years close to the end of the pre-treatment period, reducing the incidence of past shocks for the construction of the counterfactuals (Kutan and Yigit, 2007).

For the synthetic difference and difference method, we have created balanced panel data with the state as a unit of analysis from six data sets. The variable of interest collected at an individual level in the original survey was converted into the sample frame's state-level variable. The individual is the unit of analysis in the DiD method, but in synthetic DiD,

the state is the analysis unit. Other than Kerala and Karnataka, other big Indian states are also used as donor states. The panel has data for six years, from 2004 to 2017, covering six years before and after the program implementation.

This database offers comparable information across state and year. The main outcome variable of this method is the proportion/ rate of professional enrollment among the potential candidates in the state. For a stream choice analysis, the outcome variable is the proportion/rate of professional enrolment in the state among the potential candidates who have enrolled in any graduate degree program. We have used pre-program outcome data as the only predictors. The model is controlled by state-level variables such as the proportion of the potential candidates in the state, literacy rate, proportion of SC-ST population, proportion of college-going age group, the proportion of the rural population, married people proportion, and proportion of people who have completed higher secondary education in the state. We have used pre-program outcome data as the only predictors.

4.5 DiD for labour market outcomes

We used different pooled datasets to estimate the labour market outcome of the program, as explained in the data section. We aim to quantify the labour market shifts that have transpired concerning job types and sectors attributable to implementing the FGGS Scheme.

Here, we considered the potential candidates, aged 25 to 29 years and coming from no graduate household in Tamil Nadu, as a treatment group. And the same group from Kerala and Karnataka together as a control group, who would not be exposed to the FGGS. Considering the 4-6 duration for the professional degree courses in India, we assume the first batch of students exposed to the program comes to the labour market after completing the program between 2014 and 2016. Hence, we have used data collected in 2017-18 and 2018-19

to measure the post-intervention characteristics. Likewise, we have used data collected in 2009 and 2011 to measure the pre-intervention characteristics. Hence, the data from NSS 64 and 68 rounds are used as a control cohort, and PLFS 18 and 19 rounds are used as a treatment cohort.

The Difference in -Differences estimate of exposure to the FGGS program on labour market outcome is the same as Eq.1. Y_{idst} is the outcome variable, binary variable, of interest for individual, i from district, d, state, s, and time, t. We have used multiple binary variables as outcome variables to measure employment and industry types conditional on being employed. The separate outcome variable, prof_grad, is used for estimating the first-order effect. The variable 'prof_grad', which indicates whether the potential candidate has graduated with an engineering/medical/agricultural degree or not, takes the value one if the individual has a professional graduation degree; otherwise, 0. The 'post' dummy takes the value of 1 if the year is 2017 or 2018 and 0 if the year is 2009 or 2011. 'TN' is a dummy that takes the value 1 if the i-th observation comes from Tamil Nadu; otherwise, it is 0. The confounding individual, household and village characteristics that may affect the outcome variables are controlled by X_{idst} . District and year-fixed effects are used, and all the standard errors are clustered at the district level.

We also examined pre-intervention trends for all of our outcome variables using the Eq.2. The sample is confined to the pre-treatment period, with $time_t$ assuming values 1 and 2 for the years 2009 and 2011, respectively. We also added one more pre-period data in the sample, NSS round 64 for the year 2007, to examine pre-intervention trends where $time_t$ assumes values 1, 2 and 3 for the years 2004, 2009 and 2011, respectively.

5 Results

5.1 Main results

Pre-trend analysis estimations for professional course enrollment and stream choice from equation (2) are provided in Table A.2 in the appendix. The findings indicate our inability to reject the null hypothesis of parallel trends in all these specifications. This suggests that the professional course enrollment and stream choice among first-generation individuals in Tamil Nadu remain invariant over time compared to Karnataka and Kerala before the implementation of the FGGS. This ensures that the post-implementation educational gains we observe can indeed be interpreted as the causal effect of the intervention.

The results of our main regression analysis are presented in Table 2. The sample is defined as 17-22 year individuals from first-generation households of Tamilnadu, Kerala, and Karnataka in (1) - (3), and it is further restricted to those individuals who are enrolled in any graduate course in (4) – (6).

Columns (1) to (3) estimate the impact of the FGGS program on professional course enrollment Only with controls, year and state-fixed effects, and year and district-fixed effects, respectively. We find a significant positive effect of the program on professional course enrollment, indicating that the program has a substantial impact. Magnitude and statistical significance are consistent across the three models. Individuals eligible for and exposed to the FGGS scheme, first-generation individuals of Tamil Nadu in post-2010, were 3.7 percentage points more likely to enrol in professional courses than those who were not exposed.

The "Female" variable (detailed table of Table 2 in Table A.10) consistently shows a negative coefficient across all specifications, indicating that females are less likely to enrol in professional courses than their male counterparts. The coefficients for social group and

religion variables reveal disparities in enrollment patterns. Scheduled Tribe and Scheduled Caste individuals are less likely to enrol in professional courses than the General category. Additionally, Muslim individuals exhibit negative coefficients compared to Hindus.

We also observed that having an educated member in the family is positively associated with professional enrollment. The results suggest that individuals who live with household members with higher levels of prior education are more inclined to pursue professional courses, emphasizing the importance of recognizing first-generation individuals as a separate social group for affirmative action.

Columns (4) to (6) of Table 2 estimate the impact of the FGGS program on stream choice. We use the main variable of interest and the control variables in Columns (4). We add year and state fixed effects in Columns (5), year and district fixed effects in Columns (6), and control variables.

The statistically significant and positive effect shows that people moved from non-professional to professional degrees because of the FGGS scheme. The magnitude and statistical significance are consistent across the three models. Among the people who enrolled in a graduation degree, Individuals who were eligible for and exposed to the FGGS scheme were about eight percentage points more likely to shift to professional courses compared to those who were not exposed.

Similar to professional enrollment, the coefficients for gender, social group and religion variables reveal disparities in enrollment patterns. Stream choice as a professional course is less likely for women than men, SC, ST than General caste, and Muslims than Hindus. However, the magnitude of the difference is much higher in stream choice than in enrollment. For instance, the results show that females are around 12 per cent less likely than

males to enrol in professional courses. However, among those who enrolled in any graduation course, females are around 59 per cent less likely than boys to choose the professional stream.

Likewise, people from the SC category are around 17 percentage points less likely than those from the general category to enrol in professional courses. However, among those who enrolled in any graduation course, SC category individuals are around 52 percentage points less likely than boys to choose the professional stream. Furthermore, being in an urban and having household members with higher levels of prior education are positively associated with choosing a professional stream.

5.2 Robustness checks

5.2.1 Event study

The estimations of the event study analysis for professional enrollment and stream choice are given in the figure 1. These figures report coefficients from equation (2) estimation for professional course enrolment. The coefficients represent the change in enrolment for Tamilnadu relative to Kerala and Karnataka in the given years before and after the program's implementation, compared to the year 2009. A 95% confidence interval is plotted along with the coefficient. The sample is defined as 17-22-year-old individuals from first-generation households of Tamilnadu, Kerala, and Karnataka in (a), and it is further restricted to those individuals who are enrolled in any graduate course in (b). Controls include sex, marital status, relation to the head, age, social group, religion, household size, rural, number of colleges in the district, and year and state fixed effects.

Figure 1 shows the effect on professional enrollment and stream choice for each year starting from 2004 to 2017, with the base year as 2009. In both figures, the coefficients in 2004, 2007, and 2009 are statistically insignificant, indicating no significant pre-program changes

in the difference between Tamil Nadu and control states. Hence, we cannot reject the null hypothesis that the control and treated groups are statistically identical before treatment. On the other hand, we observed significant positive changes in the post-intervention period, 2011, 2014 and 2017, on both professional enrollment and stream choice. Hence, the event study supports the parallel trends assumption in DID estimation, and the significant and positive difference in the post-program intervention period can be attributed to the FGGS scheme.

5.2.2 Placebo tests

We also conducted a series of placebo run by pseudo year of implementation. In the first iteration, we restricted our data to two pre-treatment years, 2004 and 2007, and assigned 2007 as the pseudo-year of implementation of FGGS and then examined the impact of this placebo treatment using equation 2 for professional enrollment and stream choice separately. In the next iteration, the sample is restricted to 2004,2007 and 2009 and assigned 2009 as a pseudo year of implementation of FGGS; 2004 and 2007 are control years. For the years after the program's implementation, all years from 2011 are considered to be treated. Since the first graduate individuals of Tamil Nadu were not potentially exposed to the FGGS scheme between 2004 and 2009, we expect null results in the placebo-treated years. The impact on professional enrollment and stream choice should be statistically indistinguishable from zero for the placebo-treated years.

These figures, 2 a and b, report coefficients from the estimation of equation(1) for professional course enrolment (a) and stream choice (b). Each coefficient represents Tamilnadu's enrolment change relative to Kerala and Karnataka in the given year and years before that particular year. The last year is considered treated for the years before the program's implementation. After the implementation, all years from 2011 are considered to be treated.

The estimates at the points 2007 and 2009 are not significant for both professional enrollment and stream choice, i.e. the impact estimate is not significant in placebo-treated years. Also, the statistically significant and positive estimates of the post-intervention periods support our main results. Based on this placebo test and the event study analysis, we cannot reject the null hypothesis of no significant pre-treatment trends, thus adding credence to our causal inference.

5.2.3 Synthetic-DiD

Synthetic-DiD estimates of program effects are shown in Figure 3 for professional enrollment at (a) and for stream choice (b). In the DiD method, we have selected the neighbouring states of Kerala and Karnataka as the control states. Considering the socioeconomic and cultural similarities, there are chances that it can be biased and mismatched. In S-DiD, we have used a large number of states that are comparable to the treatment state, Tamil Nadu. We have used 18 states from 2004– 2017 in our donor pool.

The results show that the professional enrollment rate in the synthetic counterfactual state, where the program is not implemented, is lower than the actual trend in Tamil Nadu. Strem choice also has a lower rate in the synthetic non-exposure state. Hence, the results imply the positive impact of the FGGS Scheme. An average treatment effect on the treated (ATT) and standard errors are reported in the figure. Standard errors are based on pseudo-random placebo reshuffling, as Arkhangelsky et al. (2021) suggested for a small number of treated units.

ATT is 0.045 for professional enrollment, and ATT of professional stream choice rate is 0.076. That is, Without this program, the professional graduation enrolment rate of the potential candidates of Tamil Nadu might be approximately 4.5 percentage points lower than the ac-

tual enrolment rate of the potential candidates of Tamil Nadu; likewise, professional stream choice over non-professional courses would be 7.6 percentage points lower than the actual.

5.2.4 Border Districts

In an effort to enhance the robustness of our empirical analysis, we conducted an additional investigation by constraining the DiD estimation sample exclusively to the border districts. Specifically, we assessed the impact of the First Generation Graduate Scheme (FGGS) within the context of Tamil Nadu by isolating districts that share borders with control states, namely Kerala and Karnataka. This analysis reduces our sample to 11,959 for first-generation individuals and further reduces to 4,269 for conditioned-on graduation enrollment, encompassing a total of 26 districts: 11 districts from Tamil Nadu, seven from Karnataka, and eight from Kerala.

Table A.1 presents the results of a Difference-in-Differences (DID) analysis of the alternate sample restricted to border districts. The outcomes derived from the regression analyses exhibited qualitative consistency with the primary results. Notably, a statistically significant positive effect on professional course enrollment and the selection of professional courses persisted, mirroring the overarching trends observed in the main findings. However, the magnitude of the impact on stream choice exhibited a reduction in the border district sample.

5.3 Heterogenous effect of educational outcomes

FGGS Scheme may have a heterogeneous bearing across potential candidates in different socioeconomic groups. People are divided into social categories in India. The cultural and economic differences between them are estimated through various literature sources. Since

the tuition fee waiver is substantial and a partial amount in the overall expenses of professional courses, the program may have differential effects for male and female members of the same household, between households of different castes and religions. Urban-rural differences in terms of exposure, even the education qualification of household members, also play an essential role in awareness of availing of the program.

We analyze the heterogeneity of the DiD estimates with a modified version of equation 1, where we add the dummies of corresponding demographic and socioeconomic characteristics as the third interaction with the Post*treatment interaction term of Equation 1. For instance, we interact the gender dummy with the Post*treatment interaction term of Equation (1) to estimate the heterogeneous effects of the FGGS scheme on Individuals of different gender. The heterogeneous effect of professional enrollment is presented in figure 4, and the heterogeneous effect of stream choice is presented in figure 5.

The FGGS scheme's impact on professional enrollment is positive and significant for both males and females. However, the size of the effect is higher for males than females. Regarding stream choice, the positive and significant effect was only observed for males. Among the social categories, the effect is significant and positive only for OBC in terms of both professional enrollment and stream choice. The effect is insignificant for SC, ST and Others (General). Likewise, the effect is significant only for Hindus, not for Muslims and others.

The impact of the FGGS scheme is positive and significant for both rural and urban in terms of professional enrollment and stream choice. The difference is slightly higher for urban than rural. The program effect is not only significant for the individual who lives in a household where the education level of the older generation is above secondary. The higher effect of the FGGS scheme on Males, OBC and Urban over their counterparts implies that the program has more effect on the upper strata of the underprivileged individuals.

Professional courses were more expensive than non-professional courses. Without universal free education or the full fee waiver, professional courses cannot be accessible to all. The expenses may not be affordable for some people even after the tuition fee waiver. At the same time, the tuition fee increases affordability for those for whom the fee was partially affordable earlier.

5.4 Labour market consequences

Changes in enrollment and stream choice at the graduation level affect the labour market outcomes of the individual after graduation. We have generated another pooled household data to follow the graduated cohort of the program. Before discussing the main regression results on labour market outcomes, we must test for the first-order effect in the new pooled data. Here, we use professional graduation instead of professional enrollment. We expect the program's positive effect on the professional graduation of first-graduate individuals in Tamil Nadu.

We report the results of estimating equation 1 with district and year fixed effects in Table 3. Column (1) reports the professional graduation, and column (2) reports the stream choice. We observed a statistically significant positive effect on the likelihood of individuals graduating with professional degrees, both in the overall sample and in the sub-sample of those who have graduated. Interestingly, the magnitude of the effects, 3.4 percentage points for professional graduation and 8.5 percentage points for stream choice is also similar to the main results presented in Table 2.

Table 4 presents the results of a Difference-in-Differences (DID) analysis to evaluate the impact of the FGGS program on the employment of the individual. We have used several outcome variables derived from the primary activity of the person, such as self-employed

(Column 1), regular salaried employment (Column 2), casual employment (Column 3), seeking employment (Column 4)and Not in the labour force (Column 4). The binary variable "seeking employment" has statistically moderate and positive effects at 95% significance, and the 'casual employment" has a statistically moderate negative effect, implying that first-generation candidates employed in casual jobs are reduced and more people are looking for jobs (better jobs). Also, the results show that the number of people who were employed in regular jobs also increased by 3.7 percentage points at 90% significance. The program's impact on self-employment is not statistically significant.

This regression table 5 presents the results of a Difference-in-Differences (DID) analysis that examines the impact of the FGGS program on occupational choice among employed potential candidates. The dependent variable in the analysis assesses whether the primary activity of these individuals falls into specific occupational categories: Primary sector in column 1, Secondary sector in column 2, and Service sector in column 3. We observed that the treated group has a significant and positive association with service sector employment and a statistically moderate and negative association with the primary sector. The program's impact on the secondary sector is negative but not statistically significant. The results highlight the potential impact of FGGS in steering individuals towards service sector employment while discouraging engagement in agriculture and related sectors.

Table 6 presents regression results that estimate the impact on four distinct skill tiers (namely, Skill Level 4, Skill Level 3, Skill Level 2, and Skill Level 1) based on the National Classification of Occupation (NCO). The outcomes reveal that only Skill Level 3 exhibits a positive and statistically moderate impact (0.028). This observation implies that subsequent to the introduction of FGGS, there has been a discernible rise in the prevalence of individuals engaged in Skill Level 3 (Associate Professionals) occupations in Tamil Nadu.

The validity of the findings as causal hinges on the adherence to the parallel trends assumption, allowing us to scrutinize whether any pre-intervention trends systematically impacted the treated or control state. Pre-trend analysis estimations for the labour market outcomes from equation (2) are provided in the appendix for activity table A.5, occupational choice A.6 and occupational skills A.7. The findings indicate our inability to reject the null hypothesis of parallel trends in all these specifications. This suggests that the labour market outcomes among first-generation individuals in Tamil Nadu remain invariant over time compared to Karnataka and Kerala before the implementation of the FGGS. The results are robust across both samples: the main sample, comprising years 2009 and 2011, and the extended sample, encompassing years 2007, 2009, and 2011.

5.5 Discussion

Expenditure on professional degree education. If the FGGS scheme is driving the rise in professional enrollment among FGGs, it implies that they might be paying lower course fees. Our aim was to compare the average tuition fees, course fees, and overall education expenses for professionally enrolled undergraduate students in Tamil Nadu before and after the implementation of the program.

Since the FGGS scheme directly transfers the tuition fee, a component of the course fee, directly to colleges on behalf of students, it is expected that among students enrolled in professional courses, the out-of-pocket education expenses and course fees of the beneficiaries should be lower than those of non-beneficiaries (non-first-generation graduates). Therefore, we also compared the education expenditures between first-generation graduate (FGG) individuals and non-FGG individuals who had enrolled in professional courses within Tamil Nadu.

Unfortunately, education expenditure details were not available in all our data. Only NSSO round 64, "Participation and expenditure in Education", for the year 2007, and round 71 "" Social consumption: education", for the year 2014, provided individual-level details of education expenditure. The data included the annual course fee of the individuals currently enrolled, covering tuition fees, examination fees, development fees, and other compulsory payments. The total expenditure for the same course value included course fees, expenses on books, stationery, uniforms, transport, private coaching, and other related costs. We compare the real value of course fees and total expenditure in education in our analysis.

Figure A.1, first row, presents the average education expenditure and course fees paid by FGGs and non-FGGs before and after the FGGS scheme in Tamil Nadu. There were no significant changes between FGGs and non-FGGs in terms of education expenditure and course fees among those enrolled in professional courses in the pre-treatment period. The average education expenditure for non-FGGs was Rs 83,604, and the course fee was Rs 62,476. For FGGs, the average expenditure was Rs 86,802, and the course fee was Rs 68,500. However, after the implementation of the FGGS scheme, non-FGGs spent 43% more on education expenditure and 48% more on course fees than FGGs for professional courses. The average course fee was Rs 91,037 for non-FGGs and Rs 61,479 for FGGs.

Poorer households tend to allocate a higher share of their annual total consumption expenditure to higher education. People with past-generation graduates in their family are likely to be socioeconomically different, possibly more economically advanced, than those without graduates in the past generation. Consequently, the share of education expenditure in the overall household expenditure would be lower for the former compared to FGGs. To account for this, the share of expenditure on a professional degree for an individual is calculated as a fraction of the household's total annual expenditure, and the share of course fees is calculated similarly. A household's total annual expenditure is estimated by multiplying

its usual monthly consumption expenditure by 12 and adding the household's annual expenditure on education.

Row 2 of Figure A.1 compares the share of expenditure and share of course fees among FGGs and non-FGGs before and after the FGGS scheme in Tamil Nadu. As expected, the share of expenditure on a professional degree and the share of course fees are 32% and 29% higher, respectively, for FGGs than for non-FGGs. FGGs allocate more money from their household expenditure towards professional degree courses. After the FGGS scheme, the difference between FGGs and non-FGGs reduced to 0.04% for the share of expenditure and 0.02% for the share of course fees. The results suggest that the share of education expenditure for FGGs and non-FGGs is almost the same after implementing the FGGS scheme in Tamil Nadu. The triple difference estimate with Karnataka and Kerala as comparative states reflects the same result, indicating a decrease in educational expenditure and the share of educational expenditure among FGGs who enrolled in professional courses in Tamil Nadu after the implementation of the FGGS scheme (Table A.9).

Income and social welfare. We assess the impact of professional graduate enrollment on earnings and social welfare. However, our study encounters limitations in accessing individual-level income data, as it is only partially available in the NSS rounds utilized. Specifically, data on self-employment income is only present in a few rounds, and monthly income details are not there. However, the NSS records wage and salary earnings (received or receivable) for work done during the week, focusing on individuals employed in regular and casual jobs.

To estimate weekly earnings, we use the weekly salary for those employed in regular jobs and the weekly earnings (weekly salary + total daily wage earned in the week) for individuals engaged in both regular and casual jobs. Although the estimates lack statistical significance,

they indicate a negative trend, suggesting a decline in salary for first-generation graduates (FGGs).

In the estimation of labour market outcomes in the last session, we observed a shift among FGGs towards salaried labour and the service sector after the implementation of the FGGS. The estimation of the earnings implies that the increased labour supply may have contributed to a reduction in salaries in Tamil Nadu. However, the sample size and issues in the earning measurements do not provide sufficient strength for conclusive interpretations.

As a social welfare indicator, we examine changes in monthly per capita expenditure (MPCE) for FGGs. Did estimates in Table A.8, Column (1) present real MPCE, and Column (2) presents log MPCE. Estimates in Panel A indicate a 27% increase in MPCE among FGGs after the FGGS scheme compared to control groups. In Panel B, we find a 17% increase in MPCE for FGGs in Tamil Nadu who transitioned from a non-graduate degree to a degree course.

The estimated returns to professional graduate enrollment on earnings and social welfare suggest that, although the income is lower for professional graduates in the treatment group compared to the control group, the income is positive and higher in the unconditional samples, which include non-professional graduates (Panel 2) and both non-professional graduates and non-graduates (Panel 1). Suggesting that the changes within the FGGs MPCE are driven by the FGGS scheme.

6 Conclusion

This study evaluates the effect of the first-generation graduate scholarship (FGGS) scheme implemented in Tamil Nadu since 2010. Based on this scheme, the Tamil Nadu government pays the tuition fee of first-generation individuals enrolling in professional courses such as engineering, medicine and agriculture. We employ the exogenous timing in the program's implementation for identification and use DID regression to assess the ITT program impact on enrollment and the potentially associated benefits on labour market outcomes.

We also employed an event study, synthetic DiD, and placebo tests for robustness. Our estimates suggest a robust and significant impact of the program on enrollment. The FGGS scheme increased the professional graduate enrollment among first-generation individuals, and the scholarship made more people from the target group shift to professional courses from non-professional courses.

We also identified the heterogeneous effect of educational outcomes on first-generation individuals' demographic and socioeconomic characteristics. The higher effect of the FGGS scheme on Male, OBC and Urban over their counterparts implies that the program has more effect on upper/privileged strata of the first-generation individuals.

We also estimate the labour market consequences of the FGGS scheme. We observed that the scholarship program was steering first-generation individuals towards service-sector employment while discouraging engagement in agriculture and related sectors. We also observe a decrease in casual employment and an increase in the waiting for the job. However, the estimates are not robust.

We wish to emphasize that our current endeavour is a work in progress, and we find our-

selves in the preliminary phase of conducting a comprehensive analysis of the labour market consequences.

References

Abadie, A. (2021). Using synthetic controls: Feasibility, data requirements, and methodological aspects. *Journal of Economic Literature*, 59(2):391–425.

Arkhangelsky, D., Athey, S., Hirshberg, D. A., Imbens, G. W., and Wager, S. (2021). Synthetic difference-in-differences. *American Economic Review*, 111(12):4088–4118.

Das, U. and Sarkhel, P. (2023). Does more schooling imply improved learning? evidence from the kanyashree prakalpa in india. *Economics of Education Review*, 94:102406.

Duflo, E. (2001). Schooling and labor market consequences of school construction in indonesia: Evidence from an unusual policy experiment. *American Economic Review*, 91(4):795–813.

Hsiao, K. P. (1992). First-generation college students (eric ed351079) eric digest, november. *Office of Educational Research and Improvement*.

Kim, A. S., Choi, S., and Park, S. (2020). Heterogeneity in first-generation college students influencing academic success and adjustment to higher education. *The Social Science Journal*, 57(3):288–304.

Kutan, A. M. and Yigit, T. M. (2007). European integration, productivity growth and real convergence. *European Economic Review*, 51(6):1370–1395.

Mason, G., Williams, G., and Cranmer, S. (2009). Employability skills initiatives in higher education: What effects do they have on graduate labor market outcomes? *Education Economics*, 17(1):1–30.

Muralidharan, K. and Prakash, N. (2017). Cycling to school: Increasing secondary school enrollment for girls in india. *American Economic Journal: Applied Economics*, 9(3):321–350.

Nuñez, A., Cuccaro-Alamin, S., and Carroll, C. D. (1998). *First-Generation Students*. National Center for Education Statistics, Washington, DC.

Padrón, Y. N. (1992). The effect of strategy instruction on bilingual students' cognitive strategy use in reading. *Bilingual Research Journal*, 16(3-4):35–51.

Pascarella, E. T., Pierson, C. T., Wolniak, G. C., and Terenzini, P. T. (2004). First-generation college students: Additional evidence on college experiences and outcomes. *The Journal of Higher Education*, 75(3):249–284.

Striplin, J. J. (1999). *Facilitating Transfer for First-Generation Community College Students*.

Terenzini, P. T., Springer, L., Yaeger, P. M., Pascarella, E. T., and Nora, A. (1996). First-generation college students: Characteristics, experiences, and cognitive development. *Research in Higher Education*, 37(1):1–22.

Thayer, P. B. (2000). *Retention of Students from First Generation and Low Income Backgrounds.*

Vargas, J. H. (2004). *College Knowledge: Addressing Information Barriers to College.* College Access Services: The Education Resources Institute (TERI), Boston, MA.

Figures and Tables

Table 1: Summary Statistics

	(1) Full sample				(2) Enrolled sample			
	mean	sd	min	max	mean	sd	min	max
Enrolled in Professional Course	0.075	0.264	0	1	0.227	0.419	0	1
TN	0.402	0.490	0	1	0.450	0.497	0	1
Post	0.510	0.500	0	1	0.735	0.441	0	1
Female	0.474	0.499	0	1	0.412	0.492	0	1
Age 17 years	0.162	0.369	0	1	0.129	0.336	0	1
Age 18 years	0.202	0.401	0	1	0.262	0.440	0	1
Age 19 years	0.160	0.366	0	1	0.246	0.431	0	1
Age 20 years	0.198	0.399	0	1	0.209	0.407	0	1
Age 21 years	0.125	0.331	0	1	0.094	0.292	0	1
Age 22 years	0.153	0.360	0	1	0.060	0.238	0	1
Household head	0.003	0.054	0	1	0.000	0.018	0	1
Unmarried Child	0.771	0.420	0	1	0.906	0.292	0	1
Other relations	0.226	0.418	0	1	0.094	0.292	0	1
Scheduled Tribe	0.031	0.174	0	1	0.019	0.138	0	1
Scheduled Caste	0.181	0.385	0	1	0.158	0.365	0	1
Other Backward Class	0.630	0.483	0	1	0.659	0.474	0	1
Other social group	0.157	0.364	0	1	0.164	0.370	0	1
Hindu	0.764	0.425	0	1	0.780	0.414	0	1
Muslim	0.172	0.377	0	1	0.129	0.335	0	1
Other religions	0.065	0.246	0	1	0.091	0.288	0	1
Household Size	4.967	2.057	2	24	4.511	1.515	2	19
Rural	0.585	0.493	0	1	0.552	0.497	0	1
Urban	0.415	0.493	0	1	0.448	0.497	0	1
No. of Colleges	50.055	74.578	0	441	54.292	75.715	0	441
Proportion of Technical Colleges	0.567	0.155	0	1	0.559	0.139	0	1
Illiterate	0.143	0.350	0	1	0.077	0.266	0	1
Primary and below	0.240	0.427	0	1	0.178	0.383	0	1
Upper primary and secondary	0.456	0.498	0	1	0.510	0.500	0	1
Above secondary	0.161	0.368	0	1	0.234	0.424	0	1
Observations	35184				11699			

Notes: Data is sourced from NSS rounds 61, 64, 66, 68, 71, and 75. The above figures are for three states, Tamilnadu, Kerala, and Karnataka. The full sample is defined as 17-22 year individuals from first-generation households. The enrolled sample is further restricted to those individuals who are enrolled in any graduate course.

Table 2: Effect of FGGS scheme on professional course enrolment

	Professional course enrollment			Stream choice		
	(1)	(2)	(3)	(4)	(5)	(6)
Post × TN	0.038***	0.037***	0.037***	0.084***	0.079***	0.081***
	(0.006)	(0.006)	(0.006)	(0.021)	(0.022)	(0.021)
Observations	35184	35184	35184	11699	11699	11699
State FE	No	Yes	No	No	Yes	No
District FE	No	No	Yes	No	No	Yes
Year FE	No	Yes	Yes	No	Yes	Yes
Mean of Dep. Variable	0.031	0.031	0.031	0.143	0.143	0.143

Notes: The dependent variable is whether the individual has enrolled in a graduate degree in engineering, medicine or agriculture. Data is sourced from NSS rounds 61, 64, 66, 68, 71, and 75. The sample is defined as 17-22 year individuals from first-generation households of Tamilnadu, Kerala, and Karnataka in (1) - (3), and it is further restricted to those individuals who are enrolled in any graduate course in (4) – (6). Robust standard errors are clustered at the district level in parentheses. *** p <0.01, ** p<0.05, * p<0.1. The detailed version of the table at annexture A.10.

Table 3: First order effect on completing the graduation in professional courses

	(1)	(2)
	Professional graduates	Professional graduates
Post × TN	0.034***	0.085***
	(0.006)	(0.027)
Observations	14375	2674
District FE	Yes	Yes
Year FE	Yes	Yes
Mean of Dep. Variable	0.027	0.136

Notes: The dependent variable is whether the individual has graduated with a degree in engineering, medicine, or agriculture. Data is sourced from NSS rounds 66, 68, PLFS 17 and PLFS 18. The sample is defined as 25-29-year individuals from first-generation households of Tamilnadu, Kerala, and Karnataka in (1), and it is further restricted to those individuals who have graduated in any course in (2). Robust standard errors are clustered at the district level in parentheses. *** p <0.01, ** p<0.05, * p<0.1.

Table 4: Effect of FGGS scheme on employment

	(1)	(2)	(3)	(4)	(5)
	Self-employed	Regular employee	Casual labour	Available for 'work'	Not labour force
Post × TN	0.021	0.037*	-0.043*	0.038**	-0.038
	(0.022)	(0.020)	(0.024)	(0.015)	(0.026)
Observations	14375	14375	14375	14375	14375
District FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Mean of Dep. Variable	0.175	0.210	0.174	0.086	0.378

Notes: The dependent variables are binary variables that indicate whether the individual worked as a self/ regular salaried/ wage employee, is available for work, and is not in the labour force. Data is sourced from NSS rounds 66, 68, PLFS 17 and PLFS 18. The sample is defined as 25-29 year individuals from Tamilnadu, Kerala, and Karnataka. Controls include sex, marital status, relation to the head, age, social group, religion, household size, rural, and number of colleges in the district. Robust standard errors are clustered at the district level in parentheses. *** p <0.01, ** p<0.05, * p<0.1. The detailed version of the table at annexure A.11.

Table 5: Effect of FGGS scheme on occupational choice

	(1)	(2)	(3)
	Agriculture sector	Secondary sector	Service
Post × TN	-0.056	-0.004	0.062**
	(0.034)	(0.036)	(0.031)
Observations	8090	8090	8090
District FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Mean of Dep. Variable	0.251	0.340	0.435

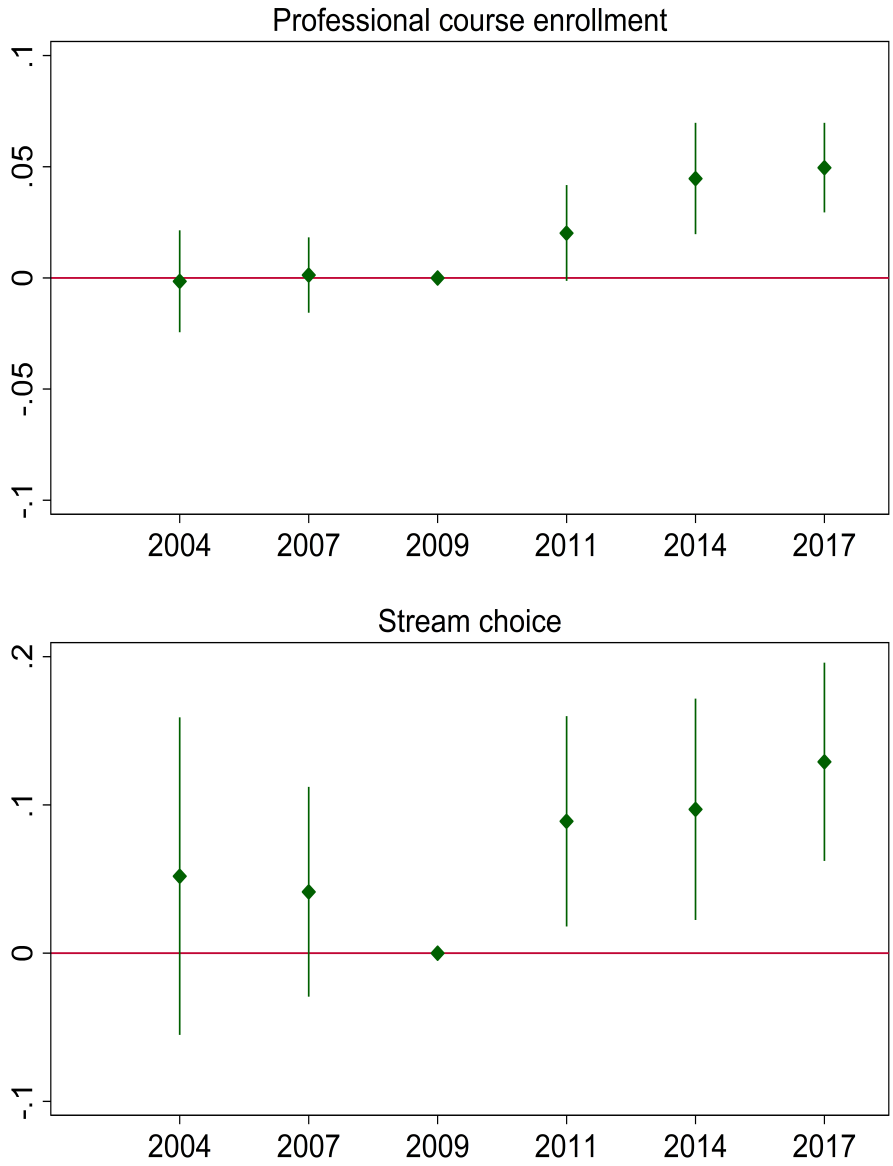
Notes: The dependent variable is whether the primary activity of the individual is in the following activities: Primary, Secondary (Manufacturing), and Service sector. Data is sourced from NSS rounds 66, 68, PLFS 17 and PLFS 18. The sample is defined as 25-29-year-old employed individuals from first-generation households of Tamilnadu, Kerala, and Karnataka (1) - (3). Controls include sex, marital status, relation to the head, age, social group, religion, household size, rural, and number of colleges in the district. Robust standard errors are clustered at the district level in parentheses.*** p <0.01, ** p<0.05, * p<0.1. The detailed version of the table at annexure A.12.

Table 6: Effect of FGGS scheme on occupational skills

	(1)	(2)	(3)	(4)
	Skill level 4	Skill level 3	Skill level 2	Skill level 1
Post × TN	0.006	0.038*	0.020	-0.029
	(0.015)	(0.022)	(0.037)	(0.038)
Observations	8090	8090	8090	8090
District FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Mean of Dep. Variable	0.061	0.068	0.532	0.237

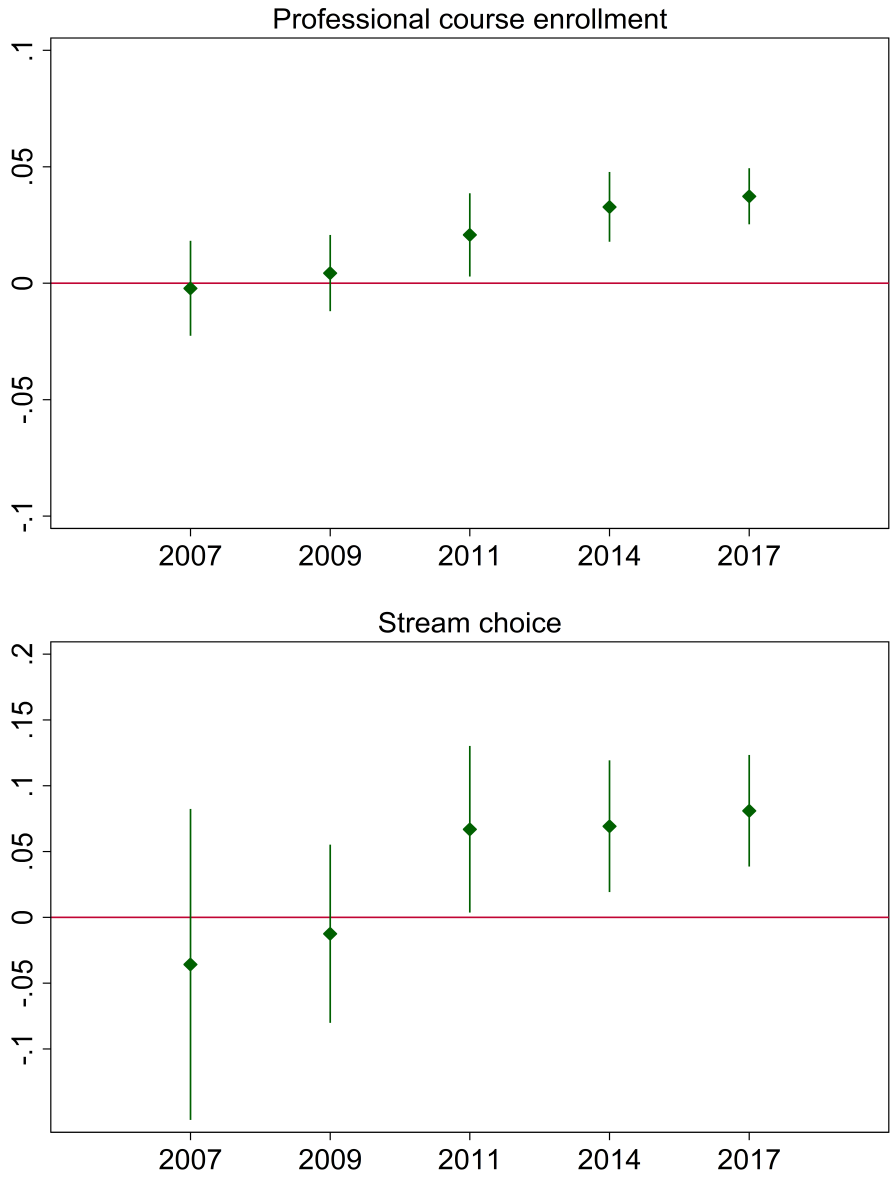
Notes: The dependent variables are binary variables to indicate the skill level of the individuals based on NCO 2004. Data is sourced from NSS rounds 66, 68, PLFS 17 and PLFS 18. The sample is defined as 25-29-year-old employed individuals from first-generation households in Tamilnadu, Kerala, and Karnataka (1) - (4). Controls include sex, marital status, relation to the head, age, social group, religion, household size, rural, and number of colleges in the district. Robust standard errors are clustered at the district level in parentheses.*** p <0.01, ** p<0.05, * p<0.1. The detailed version of the table at annexure A.13.

Figure 1: Event study analysis for professional enrolment



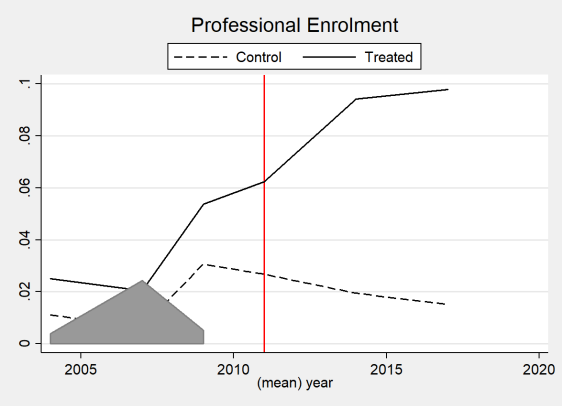
Notes: These figures report coefficients from the estimation of equation (2) for professional Course enrolment. The coefficients represent the change in enrolment for Tamilnadu relative to Kerala and Karnataka in the given years before and after the implementation of the program, as compared to the year 2009. 95% confidence interval is plotted along with the coefficient. The sample is defined as 17-22 year individuals from first-generation households of Tamilnadu, Kerala, and Karnataka in (a), and it is further restricted to those individuals who are enrolled in any graduate course in (b). Controls include sex, marital status, relation to the head, age, social group, religion, household size, rural, number of colleges in the district, and year and district fixed effects.

Figure 2: Placebo Test for professional enrolment

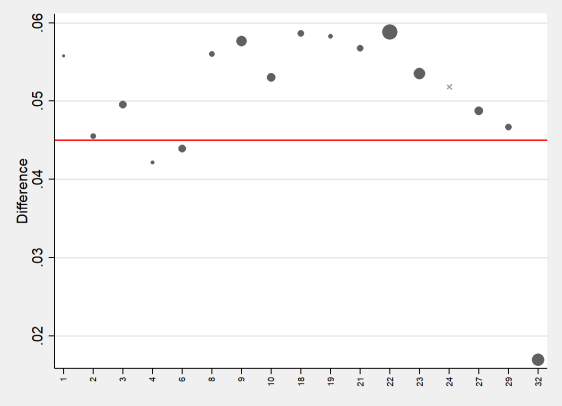


Notes: These figures report coefficients from the estimation of DID equation for prof. course enrolment. Each coefficient represents the change in enrolment for Tamilnadu relative to Kerala and Karnataka in the given year and years before that particular year. For the years before the implementation of the program, the last year is considered as treated. After the implementation, all years from 2011 is considered to be treated. For instance, the coefficient corresponding to 2009 represents the DID coefficient with data from 2004, 2007, and 2009, and 2009 is considered to be treated. 95% confidence interval is plotted along with the coefficient. The sample is defined as 17-22 year individuals from first-generation households of Tamilnadu, Kerala, and Karnataka in (a), and it is further restricted to those individuals who are enrolled in any graduate course in (b). Controls include sex, marital status, relation to head, age, social group, religion, household size, rural, number of colleges in the district, and year and district fixed effects.

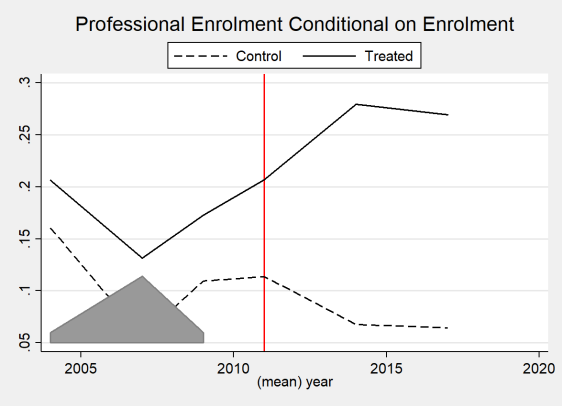
Figure 3: Effect of FGGS scheme on professional enrolment: Synthetic DiD



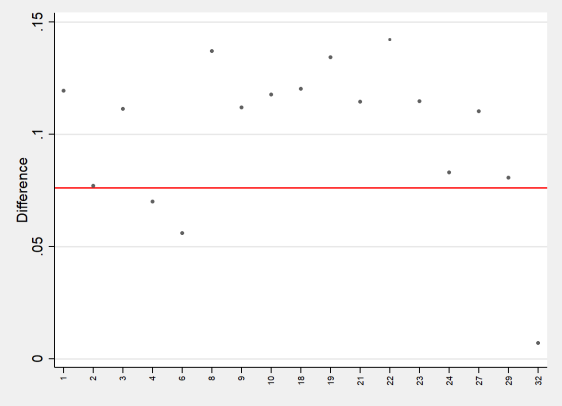
(a) All sample: ATT 0.045, se 0.0057



(b) Weights



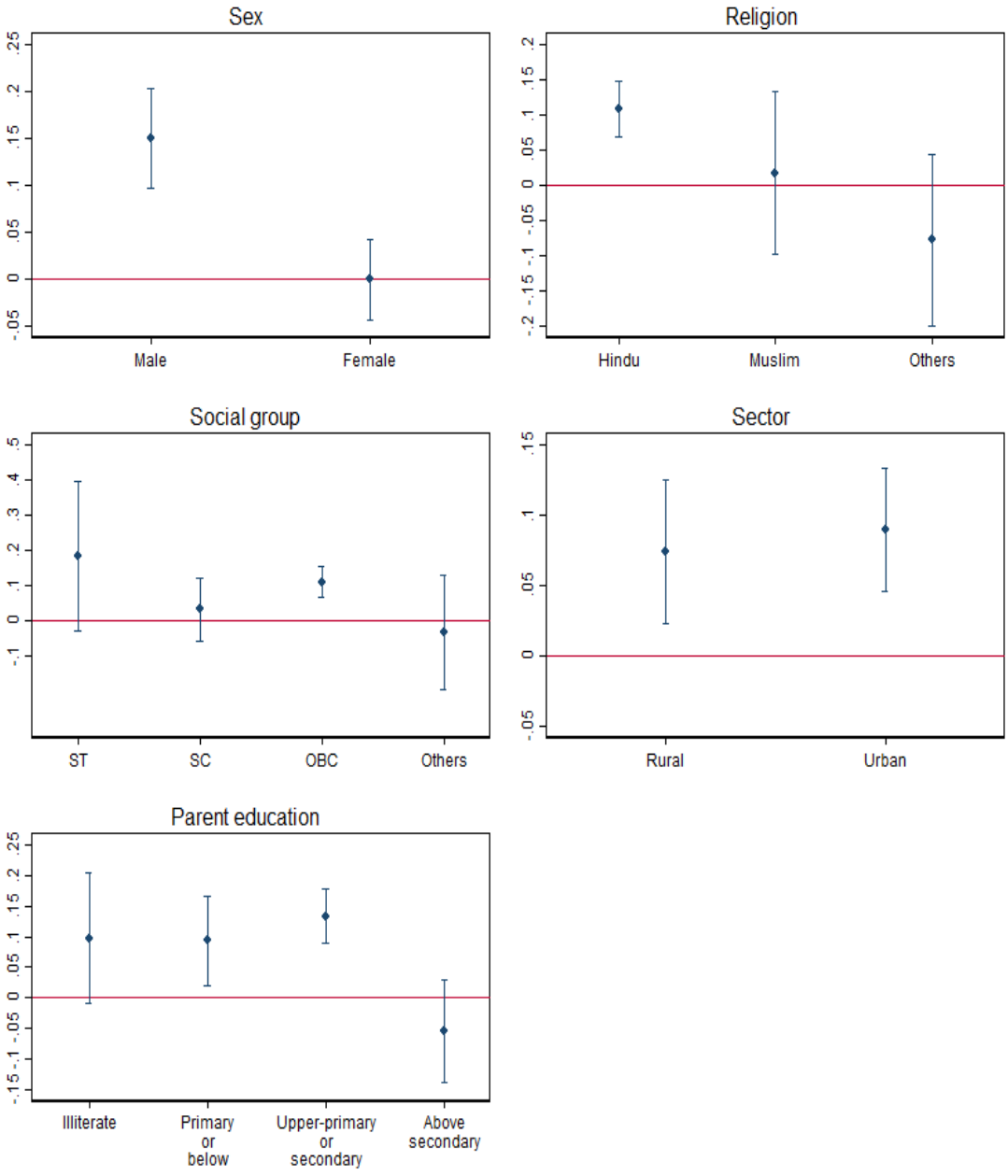
(c) Enrolled sample: ATT 0.076, se 0.025



(d) Weights

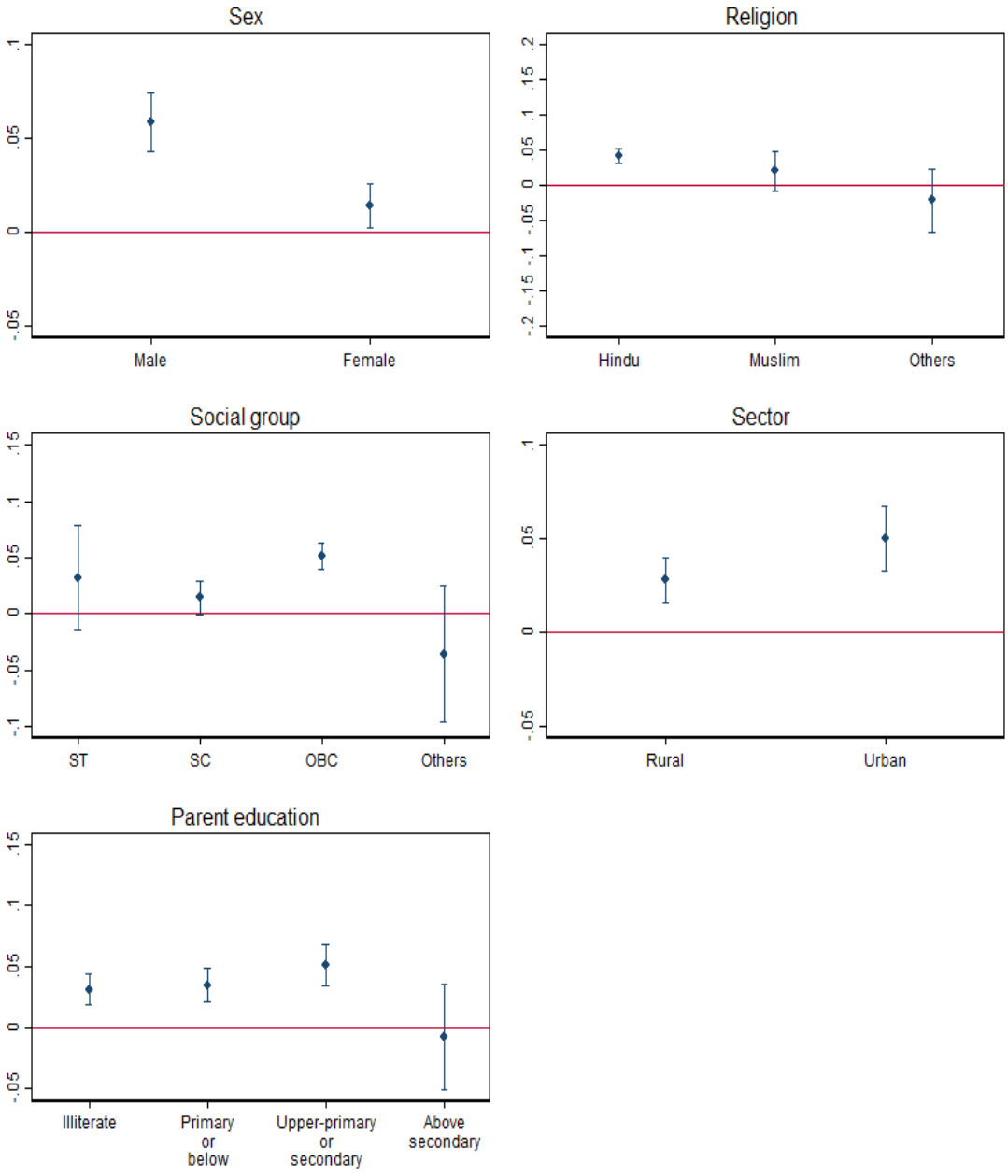
Notes: These figures are generated using the synthetic difference in differences method for professional course enrolment. The variable is the proportion of individuals enrolled in professional courses from all 17-22 year individuals from first-generation households in (a), and it is individuals enrolled in professional courses from 17-22 year individuals enrolled in any graduate course from first-generation households in (c).

Figure 4: Heterogeneity in the effect of FGGS on professional enrollment by sex, religion, social group, sector and parent education



Notes: The above figure plots the DiD estimates of the FGGS on professional enrollment on different sub-samples. We consider different sub-samples based on (a) sex, (b) Religion, (c) Social group, (d) Sector, and (e) Parent education. All regressions control for district and year-fixed effects and other controls.

Figure 5: Heterogeneity in the effect of FGGS on stream choice by sex, religion, social group, sector and parent education



Notes: The above figure plots the DiD estimates of the FGGS on stream choice on different sub-samples.We consider different sub-samples based on (a) sex, (b) Religion, (c) Social group, (d) Sector, and (e) Parent education. All regressions control for district and year-fixed effects and other controls.

Appendix.

Table A.1: Effect on professional course enrolment (Border districts sample)

	Professional course enrollment			Stream choice		
	(1)	(2)	(3)	(4)	(5)	(6)
Post × TN	0.0385***	0.0375***	0.0386***	0.0573**	0.0504*	0.0522*
	(3.38)	(3.31)	(3.62)	(2.29)	(2.03)	(1.95)
Observations	11959	11959	11959	4269	4269	4269
State FE	No	Yes	No	No	Yes	No
District FE	No	No	Yes	No	No	Yes
Year FE	No	Yes	Yes	No	Yes	Yes
Mean of Dep. Variable	0.035	0.035	0.035	0.144	0.144	0.144

Notes: Data is sourced from NSS rounds 61, 64, 66, 68, 71, and 75. The dependent variable is enrollment in Professional courses. The sample is restricted to border districts of Tamil Nadu and districts of Kerala and Karnata, which share a border with Tamil Nadu. The full sample is defined as 17-22 year individuals from first-generation households (columns 1, 2, 3). The enrolled sample is further restricted to those individuals who are enrolled in any graduate course (columns 4,5,6). Robust standard errors are clustered at the district level in parentheses. *** p <0.01, ** p<0.05, * p<0.1.

Table A.2: Pre-Trend analysis for the effects on professional enrollment

	Professional course enrollment			Stream choice		
	(1)	(2)	(3)	(4)	(5)	(6)
TN × time	0.002	0.001	0.001	−0.023	−0.025	−0.014
	(0.005)	(0.005)	(0.005)	(0.024)	(0.024)	(0.025)
Observations	17245	17245	17245	3096	3096	3096
State FE	*No*	*Yes*	*No*	*No*	*Yes*	*No*
District FE	*No*	*No*	*Yes*	*No*	*No*	*Yes*
Year FE	*No*	*Yes*	*Yes*	*No*	*Yes*	*Yes*
Mean of Dep. Variable	0.018	0.018	0.018	0.106	0.106	0.106

Notes: Data is sourced from NSS rounds 61, 64, and 66. The dependent variable is enrollment in Professional courses. The full sample is defined as 17-22 year individuals from first-generation households. The enrolled sample is further restricted to those individuals who are enrolled in any graduate course. Time takes 1, 2, and 3 for 2004, 2007 and 2011, respectively. Robust standard errors are clustered at the district level in parentheses. *** p <0.01, ** p<0.05, * p<0.1.

Table A.3: Summary Statistics - Full sample and Graduated Sample

	(1) Full sample				(2) Graduated sample			
	mean	sd	min	max	mean	sd	min	max
Regular employee	0.195	0.396	0	1	0.395	0.489	0	1
Self-employed	0.200	0.400	0	1	0.119	0.324	0	1
Casual labour	0.183	0.387	0	1	0.032	0.176	0	1
Available for 'work'	0.078	0.269	0	1	0.205	0.404	0	1
Not labour force	0.376	0.485	0	1	0.276	0.447	0	1
Skill level 4	0.034	0.180	0	1	0.130	0.337	0	1
Skill level 3	0.036	0.187	0	1	0.109	0.312	0	1
Skill level 2	0.330	0.470	0	1	0.212	0.409	0	1
Skill level 1	0.111	0.314	0	1	0.020	0.139	0	1
Agriculture sector	0.132	0.339	0	1	0.040	0.197	0	1
Secondary sector	0.195	0.396	0	1	0.092	0.289	0	1
Service	0.256	0.436	0	1	0.415	0.493	0	1
TN	0.399	0.490	0	1	0.473	0.499	0	1
Post	0.466	0.499	0	1	0.620	0.486	0	1
Professional graduates	0.018	0.134	0	1	0.098	0.297	0	1
Female	0.540	0.498	0	1	0.512	0.500	0	1
Age 25 years	0.234	0.423	0	1	0.261	0.439	0	1
Age 26 years	0.207	0.405	0	1	0.228	0.419	0	1
Age 27 years	0.179	0.383	0	1	0.187	0.390	0	1
Age 28 years	0.236	0.424	0	1	0.192	0.394	0	1
Age 29 years	0.145	0.352	0	1	0.132	0.339	0	1
Household head	0.037	0.188	0	1	0.024	0.154	0	1
Unmarried Child	0.313	0.464	0	1	0.488	0.500	0	1
Other relations	0.650	0.477	0	1	0.488	0.500	0	1
Scheduled Caste	0.031	0.173	0	1	0.014	0.117	0	1
Scheduled Tribe	0.171	0.377	0	1	0.144	0.351	0	1
Other Backward Class	0.658	0.474	0	1	0.662	0.473	0	1
Other social group	0.140	0.347	0	1	0.180	0.384	0	1
Hindu	0.767	0.423	0	1	0.801	0.399	0	1
Muslim	0.164	0.370	0	1	0.098	0.297	0	1
Other religions	0.069	0.254	0	1	0.101	0.301	0	1
Household Size	5.154	2.410	2	35	4.589	1.893	2	35
Ever married	0.651	0.477	0	1	0.472	0.499	0	1
Urban	0.455	0.498	0	1	0.573	0.495	0	1
Illiterate	0.171	0.377	0	1	0.091	0.287	0	1
Primary and below	0.234	0.424	0	1	0.159	0.366	0	1
Upper primary and secondary	0.456	0.498	0	1	0.499	0.500	0	1
Above secondary	0.138	0.345	0	1	0.252	0.434	0	1
Observations	14375				2674			

Notes: Data is sourced from NSS rounds 66, 68, and PLFS years 17-18 and 18-19. The above figures are for three states: Tamil Nadu, Kerala, and Karnataka. The full sample is defined as 15-29-year-old individuals from first-generation households. The sample is further restricted to those individuals who have graduated in any course (Graduated).

Table A.4: Pre-Trend analysis for the first order effect on graduation

	(1)	(2)
	Professional graduates	Professional graduates
TN × time	-0.009	-0.077
	(0.005)	(0.041)
Observations	7670	1014
District FE	Yes	Yes
Year FE	Yes	Yes
Mean of Dep. Variable	0.005	0.049

Notes: The dependent variables are binary variables that indicate whether the individual has a graduate degree in a professional course. Data is sourced from NSS rounds 66 and 68. The sample is defined as 25-29-year individuals from first-generation households of Tamilnadu, Kerala, and Karnataka (1).The sample is further restricted to those individuals who have graduated in any course (Graduated). Controls include sex, marital status, relation to the head, age, social group, religion, household size, rural, and number of colleges in the district. Variable "Time" takes the value of 1 2 for the years 2009 and 2011, respectively. Robust standard errors are clustered at the district level in parentheses.*** p <0.01, ** p<0.05, * p<0.1

Table A.5: Pre-Trend analysis for the effects on employment

	(1)	(2)	(3)	(4)	(5)
	Self-employed	Regular employee	Casual labour	Available for 'work'	Not labour force
TN × time	0.004	-0.004	0.029	0.002	0.011
	(0.022)	(0.009)	(0.020)	(0.008)	(0.018)
Observations	11692	11692	11692	11692	11692
District FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Mean of Dep. Variable	0.246	0.136	0.265	0.041	0.367

Notes: The dependent variables are binary variables that indicate whether the individual worked as a self/ regular salaried/ wage employee, is available for work, and is not in the labour force. Data is sourced from NSS rounds 64, 66 and 68. Note, here we have included one more year (2007/round 64) to estimate the pre-trend. Effects are also similar for the main sample, with years 2009 and 2011. The sample is defined as 25-29-year-old individuals from first-generation households in Tamil Nadu, Kerala, and Karnataka. Controls include sex, marital status, relation to the head, age, social group, religion, household size, rural, and number of colleges in the district. Variable "Time" takes 1, 2, and 3 for 2007, 2009, and 2011, respectively. Robust standard errors are clustered at the district level in parentheses.*** p <0.01, ** p<0.05, * p<0.1.

Table A.6: Pre-Trend analysis for the effects on occupational choice

	(1)	(2)	(3)
	Agriculture sector	Secondary sector	Service
TN × time	-0.030	0.049*	0.000
	(0.022)	(0.025)	(0.019)
Observations	6962	6962	6962
District FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Mean of Dep. Variable	0.412	0.301	0.329

Notes: The dependent variable is whether the primary activity of the individual is in the following activities: Primary, Secondary (Manufacturing), and Service sector. Data is sourced from NSS rounds 64, 66 and 68. Note, here we have included one more year (2007/round 64) to estimate the pre-trend. Effects are also similar for the main sample, with years 2009 and 2011. The sample is defined as 25-29-year-old employed individuals from first-generation households in Tamil Nadu, Kerala, and Karnataka. Controls include sex, marital status, relation to the head, age, social group, religion, household size, rural, and number of colleges in the district. Variable "Time" takes 1, 2, and 3 for 2007, 2009, and 2011, respectively. Robust standard errors are clustered at the district level in parentheses.*** p <0.01, ** p<0.05, * p<0.1.

Table A.7: Pre-Trend analysis for the effects on occupational skills

	(1)	(2)	(3)	(4)
	Skill level 4	Skill level 3	Skill level 2	Skill level 1
TN × time	0.010	0.002	0.048*	0.015
	(0.009)	(0.006)	(0.026)	(0.027)
Observations	6962	6962	6962	6962
District FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Mean of Dep. Variable	0.036	0.035	0.586	0.308

Notes: The dependent variables are binary variables to indicate the skill level of the individuals based on NCO 2004. Data is sourced from NSS rounds 64, 66 and 68. Note, here we have included one more year (2007/round 64) to estimate the pre-trend. Effects are also similar for the main sample, with years 2009 and 2011. The sample is defined as 25-29-year-old employed individuals from first-generation households in Tamil Nadu, Kerala, and Karnataka (1). Controls include sex, marital status, relation to the head, age, social group, religion, household size, rural, and number of colleges in the district. Variable "Time" takes 1, 2, and 3 for 2007, 2009, and 2011, respectively. Robust standard errors are clustered at the district level in parentheses.*** p <0.01, ** p<0.05, * p<0.1.

Table A.8: Effect of FGGS scheme on social welfare

	(1) MPCE(Real)	(2) MPCE
Panel A - Potential candidates		
Post=1 × TN=1	412.586***	0.278***
	(52.966)	(0.030)
Observations	14375	14375
District FE	Yes	Yes
Year FE	Yes	Yes
Mean of Dep. Variable	1425.092	7.112
	(1)	(2)
Panel B - Conditional on Graduation		
Post=1 × TN=1	500.684***	0.163***
	(166.189)	(0.055)
Observations	2674	2674
District FE	Yes	Yes
Year FE	Yes	Yes
Mean of Dep. Variable	1933.685	7.427

Notes: The dependent variable is monthly per capita expenditure (real) and log MPCE (2). Data is sourced from NSS rounds 66, 68, PLFS 17 and PLFS 18. The sample is defined as 25-29-year-old individuals from first-generation households in Tamil Nadu, Kerala, and Karnataka. Controls include sex, marital status, relation to the head, age, social group, religion, household size, rural, and number of colleges in the district. Robust standard errors are clustered at the district level in parentheses. *** p <0.01, ** p<0.05, * p<0.1.

Table A.9: Expenditure on professional degree education: DDD estimation

	(1) Education expenditure	(2) Education expenditure share	(3) Course fee	(4) course fee share
Panel A				
Post × TN × FGG	-71549.763**	-17.205*	-55692.877*	-15.450*
	(31381.546)	(8.687)	(28460.043)	(8.132)
Observations	965	965	959	959
State FE	No	No	No	No
Year FE	No	No	No	No
Mean of Dep. Variable	93783.344	31.510	69279.845	22.871
	(1)	(2)	(3)	(4)
Panel B: with year and State FE				
Post × TN × FGG	-66126.036**	-17.448*	-51268.420*	-15.656*
	(31093.552)	(8.780)	(28761.527)	(8.205)
Observations	965	965	959	959
State FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Mean of Dep. Variable	93783.344	31.510	69279.845	22.871

Notes: Data is sourced from NSS rounds 64 & 71. The full sample is defined as 17-22-year-old individuals who are enrolled in professional graduate courses. The dependent variable is educational expenditure(Real), course fee (real), the share of expenditure on the professional degree of an individual as a fraction of a household's total annual expenditure and the share of the course fee as a fraction of a household's total annual expenditure. A household's total annual expenditure is estimated by multiplying its usual monthly consumption expenditure by 12 and adding to this figure the household's annual expenditure on education. Robust standard errors are clustered at the district level in parentheses. *** p <0.01, ** p<0.05, * p<0.1.

Table A.10: Effect of FGGS scheme on professional course enrolment

	Professional course enrollment			Stream choice		
	(1)	(2)	(3)	(4)	(5)	(6)
Post × TN	0.038***	0.037***	0.037***	0.084***	0.079***	0.081***
	(0.006)	(0.006)	(0.006)	(0.021)	(0.022)	(0.021)
Post=1	0.009***			0.026**		
	(0.003)			(0.012)		
TN=1	0.007**			0.046***		
	(0.003)			(0.016)		
Female	-0.012***	-0.012***	-0.012***	-0.053***	-0.058***	-0.059***
	(0.003)	(0.003)	(0.003)	(0.010)	(0.010)	(0.011)
Age 18 years	0.020***	0.021***	0.021***	0.005	0.008	0.007
	(0.003)	(0.003)	(0.003)	(0.016)	(0.016)	(0.016)
Age 19 years	0.030***	0.031***	0.031***	0.017	0.018	0.019
	(0.005)	(0.005)	(0.005)	(0.016)	(0.016)	(0.016)
Age 20 years	0.019***	0.020***	0.020***	0.035**	0.037**	0.036**
	(0.004)	(0.004)	(0.004)	(0.017)	(0.016)	(0.017)
Age 21 years	0.011**	0.011**	0.012**	0.106***	0.103***	0.104***
	(0.005)	(0.005)	(0.005)	(0.025)	(0.024)	(0.024)
Age 22 years	-0.003	-0.002	-0.003	0.057**	0.056**	0.054*
	(0.004)	(0.004)	(0.004)	(0.027)	(0.027)	(0.028)
Unmarried Child	0.026***	0.022***	0.020***	-0.109	-0.086	-0.081
	(0.006)	(0.008)	(0.007)	(0.232)	(0.246)	(0.256)
Other relations	0.007	0.003	0.001	-0.118	-0.093	-0.086
	(0.007)	(0.008)	(0.008)	(0.237)	(0.252)	(0.262)
Scheduled Tribe	-0.014***	-0.013***	-0.013***	-0.063**	-0.056**	-0.050*
	(0.004)	(0.004)	(0.005)	(0.024)	(0.025)	(0.025)
Scheduled Caste	-0.016***	-0.017***	-0.017***	-0.049***	-0.047***	-0.052***
	(0.005)	(0.005)	(0.005)	(0.018)	(0.018)	(0.018)
Other Backward Class	-0.001	-0.002	-0.003	-0.004	-0.002	-0.009
	(0.003)	(0.003)	(0.004)	(0.012)	(0.012)	(0.012)
Muslim	-0.012***	-0.015***	-0.013***	-0.030	-0.033*	-0.025
	(0.004)	(0.004)	(0.004)	(0.018)	(0.019)	(0.020)
Other religions	0.016*	0.011	0.005	0.009	-0.000	-0.006
	(0.008)	(0.008)	(0.008)	(0.020)	(0.020)	(0.021)
Primary and below	0.005*	0.004	0.004	0.017	0.014	0.019
	(0.003)	(0.003)	(0.003)	(0.018)	(0.019)	(0.020)
Upper primary and secondary	0.025***	0.022***	0.023***	0.052***	0.049***	0.053***
	(0.004)	(0.004)	(0.004)	(0.017)	(0.018)	(0.018)
Above secondary	0.059***	0.056***	0.056***	0.103***	0.098***	0.101***
	(0.006)	(0.006)	(0.006)	(0.019)	(0.020)	(0.020)
Household Size	-0.002***	-0.002***	-0.002***	-0.007**	-0.007**	-0.007**
	(0.001)	(0.001)	(0.001)	(0.003)	(0.003)	(0.003)
No. of Colleges	-0.000	0.000	0.000**	-0.000	0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Proportion of Technical Colleges	0.022**	0.017*	0.045**	0.137***	0.115***	0.318***
	(0.009)	(0.009)	(0.020)	(0.040)	(0.041)	(0.116)
Urban	0.014***	0.014***	0.014***	0.033***	0.030***	0.030**
	(0.003)	(0.003)	(0.004)	(0.011)	(0.011)	(0.013)
Constant	-0.031***	-0.016*	-0.040**	0.095	0.122	0.013
	(0.007)	(0.008)	(0.016)	(0.233)	(0.248)	(0.268)
Observations	35184	35184	35184	11699	11699	11699
State FE	No	Yes	No	No	Yes	No
District FE	No	No	Yes	No	No	Yes
Year FE	No	Yes	Yes	No	Yes	Yes
Mean of Dep. Variable	0.031	0.031	0.031	0.143	0.143	0.143

Notes: This is the detailed version of Table 2. Robust standard errors are clustered at the district level in parentheses. *** p <0.01, ** p<0.05, * p<0.1.

Table A.11: Effect of FGGS scheme on employment

	(1) Self-employed	(2) Regular employee	(3) Casual labour	(4) Available for 'work'	(5) Not labour force
Post × TN	0.021	0.037*	-0.043*	0.038**	-0.038
	(0.022)	(0.020)	(0.024)	(0.015)	(0.026)
Female	-0.218***	-0.168***	-0.165***	0.013	0.546***
	(0.020)	(0.016)	(0.019)	(0.016)	(0.018)
Age 26 years	0.031	0.032**	-0.010	-0.019	-0.027*
	(0.020)	(0.016)	(0.019)	(0.013)	(0.015)
Age 27 years	0.040**	0.048**	0.005	-0.032**	-0.047***
	(0.017)	(0.019)	(0.018)	(0.013)	(0.017)
Age 28 years	0.033**	0.033**	0.035	-0.031**	-0.049***
	(0.017)	(0.016)	(0.021)	(0.012)	(0.017)
Age 29 years	0.078***	0.015	0.025	-0.028*	-0.076***
	(0.017)	(0.016)	(0.022)	(0.015)	(0.018)
Unmarried Child	-0.057	-0.036	-0.088	0.062*	0.114***
	(0.043)	(0.054)	(0.061)	(0.034)	(0.040)
Other relations	-0.075**	-0.029	-0.020	0.011	0.090***
	(0.029)	(0.036)	(0.034)	(0.012)	(0.025)
Scheduled Caste	-0.063*	-0.081***	0.071	0.024	0.062
	(0.035)	(0.029)	(0.047)	(0.020)	(0.038)
Scheduled Tribe	-0.121***	-0.033*	0.121***	0.026**	-0.001
	(0.022)	(0.020)	(0.029)	(0.013)	(0.017)
Other Backward Class	-0.042**	-0.015	0.017	0.010	0.014
	(0.020)	(0.017)	(0.021)	(0.011)	(0.018)
Muslim	-0.023	-0.068***	-0.050**	0.004	0.131***
	(0.019)	(0.018)	(0.019)	(0.014)	(0.015)
Other religions	-0.051***	-0.005	-0.000	0.038*	0.025
	(0.017)	(0.027)	(0.019)	(0.022)	(0.023)
Primary and below	0.029	0.007	-0.086***	0.001	0.032*
	(0.020)	(0.019)	(0.025)	(0.011)	(0.019)
Upper primary and secondary	0.032*	-0.015	-0.136***	0.035***	0.078***
	(0.016)	(0.016)	(0.022)	(0.012)	(0.020)
Above secondary	0.004	-0.005	-0.174***	0.043***	0.115***
	(0.017)	(0.025)	(0.023)	(0.015)	(0.024)
Household Size	0.005	-0.001	-0.004	0.001	-0.002
	(0.003)	(0.003)	(0.003)	(0.002)	(0.003)
Ever married	0.073**	-0.115***	-0.014	-0.071*	0.138***
	(0.035)	(0.039)	(0.049)	(0.039)	(0.037)
Rural	0.049***	-0.095***	0.044***	0.008	0.011
	(0.011)	(0.016)	(0.014)	(0.009)	(0.014)
Constant	0.256***	0.465***	0.378***	0.059*	-0.135***
	(0.048)	(0.063)	(0.064)	(0.035)	(0.044)
Observations	14375	14375	14375	14375	14375
District FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Mean of Dep. Variable	0.175	0.210	0.174	0.086	0.378

Notes: This is the detailed version of Table 4. Robust standard errors are clustered at the district level in parentheses. *** p <0.01, ** p<0.05, * p<0.1.

Table A.12: Effect of FGGS scheme on occupational choice

	(1) Agriculture sector	(2) Secondary sector	(3) Service
Post × TN	-0.056	-0.004	0.062**
	(0.034)	(0.036)	(0.031)
Female	0.108***	-0.046*	-0.056**
	(0.025)	(0.027)	(0.027)
Age 26 years	-0.020	0.023	0.008
	(0.028)	(0.028)	(0.037)
Age 27 years	-0.005	0.002	0.001
	(0.022)	(0.030)	(0.033)
Age 28 years	0.056**	-0.034	-0.014
	(0.026)	(0.025)	(0.027)
Age 29 years	0.015	0.012	-0.015
	(0.025)	(0.034)	(0.033)
Unmarried Child	-0.014	-0.023	0.012
	(0.052)	(0.040)	(0.054)
Other relations	-0.069	0.043	-0.001
	(0.044)	(0.038)	(0.039)
Scheduled Caste	0.171**	-0.084	-0.104
	(0.068)	(0.054)	(0.063)
Scheduled Tribe	-0.015	0.105***	-0.103***
	(0.033)	(0.032)	(0.032)
Other Backward Class	-0.013	0.053**	-0.068***
	(0.024)	(0.025)	(0.024)
Muslim	-0.109***	0.015	0.091***
	(0.028)	(0.034)	(0.027)
Other religions	-0.044*	0.012	0.021
	(0.025)	(0.043)	(0.040)
Primary and below	-0.008	-0.025	0.012
	(0.032)	(0.026)	(0.028)
Upper primary and secondary	-0.031	-0.021	0.043
	(0.029)	(0.025)	(0.028)
Above secondary	-0.055*	-0.069**	0.110***
	(0.031)	(0.034)	(0.034)
Household Size	0.002	-0.003	0.002
	(0.004)	(0.005)	(0.005)
Ever married	0.076*	-0.026	-0.039
	(0.044)	(0.046)	(0.041)
Rural	0.233***	-0.020	-0.198***
	(0.029)	(0.029)	(0.024)
Constant	0.113*	0.349***	0.593***
	(0.066)	(0.058)	(0.068)
Observations	8090	8090	8090
District FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Mean of Dep. Variable	0.251	0.340	0.435

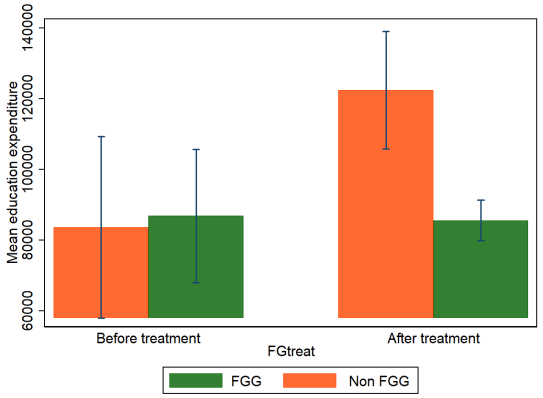
Notes: This is the detailed version of Table 5. Robust standard errors are clustered at the district level in parentheses. *** p <0.01, ** p<0.05, * p<0.1.

Table A.13: Effect of FGGS scheme on occupational skills

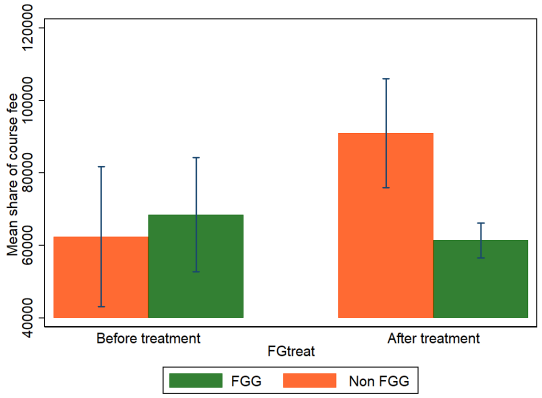
	(1) Skill level 4	(2) Skill level 3	(3) Skill level 2	(4) Skill level 1
Post × TN	0.006	0.038*	0.020	-0.029
	(0.015)	(0.022)	(0.037)	(0.038)
Female	0.047***	0.055***	-0.159***	0.121***
	(0.014)	(0.012)	(0.031)	(0.027)
Age 26 years	-0.005	0.010	0.024	-0.021
	(0.012)	(0.012)	(0.029)	(0.026)
Age 27 years	-0.000	-0.014	0.008	-0.030
	(0.014)	(0.013)	(0.036)	(0.026)
Age 28 years	-0.012	-0.023**	0.008	0.023
	(0.012)	(0.010)	(0.029)	(0.029)
Age 29 years	0.001	-0.009	-0.026	-0.000
	(0.016)	(0.013)	(0.035)	(0.028)
Unmarried Child	0.016	-0.009	0.086	-0.084
	(0.024)	(0.026)	(0.069)	(0.068)
Other relations	-0.024	-0.007	0.053	-0.034
	(0.023)	(0.016)	(0.044)	(0.041)
Scheduled Caste	-0.050**	-0.001	-0.043	0.138*
	(0.020)	(0.023)	(0.065)	(0.071)
Scheduled Tribe	-0.035*	-0.022	-0.083*	0.141***
	(0.019)	(0.016)	(0.045)	(0.035)
Other Backward Class	-0.012	-0.007	-0.008	0.017
	(0.017)	(0.016)	(0.037)	(0.032)
Muslim	-0.042***	-0.001	-0.091***	0.015
	(0.012)	(0.013)	(0.034)	(0.025)
Other religions	0.001	0.019	-0.045	-0.010
	(0.023)	(0.030)	(0.043)	(0.028)
Primary and below	0.008	-0.016	0.050	-0.094***
	(0.012)	(0.014)	(0.037)	(0.034)
Upper primary and secondary	0.029***	0.008	-0.015	-0.119***
	(0.011)	(0.014)	(0.028)	(0.028)
Above secondary	0.100***	0.034	-0.091**	-0.147***
	(0.028)	(0.029)	(0.041)	(0.041)
Household Size	-0.003	0.002	0.008	0.000
	(0.003)	(0.002)	(0.006)	(0.005)
Ever married	0.020	-0.040	-0.008	0.022
	(0.013)	(0.025)	(0.048)	(0.045)
Rural	-0.043***	-0.037**	0.023	0.117***
	(0.009)	(0.017)	(0.023)	(0.019)
Constant	0.086**	0.094***	0.492***	0.219***
	(0.033)	(0.035)	(0.069)	(0.082)
Observations	8090	8090	8090	8090
District FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Mean of Dep. Variable	0.061	0.068	0.532	0.237

Notes: This is the detailed version of Table 6. Robust standard errors are clustered at the district level in parentheses. *** p <0.01, ** p<0.05, * p<0.1.

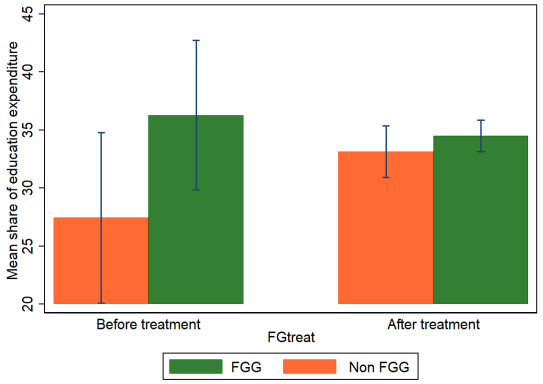
Figure A.1: Expenditure on professional degree education



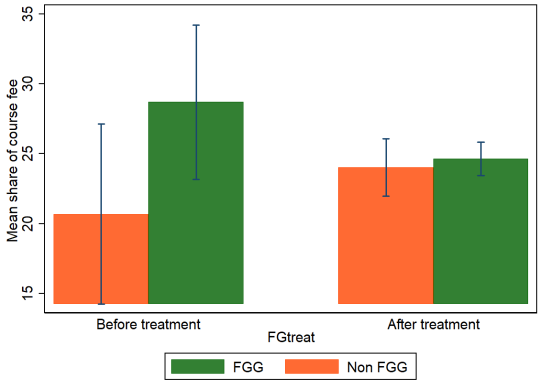
(a) Expenditure in education



(b) Course fee



(c) Share of expenditure in education



(d) Share of course fee

Notes: Data is sourced from NSS rounds 64 & 71. The full sample is defined as 17-22-year-old individuals who are enrolled in professional graduate courses. The dependent variable is educational expenditure(Real), course fee (real), the share of expenditure on the professional degree of an individual as a fraction of a household's total annual expenditure and the share of the course fee as a fraction of a household's total annual expenditure.