Gender gap in schooling: Is there a role for health insurance?*

Manini Ojha†

November, 2017

Abstract

Health shocks can have significant consequences for human capital of future generations in countries with a poor system of health insurance. Access to health insurance may not only play a role in determining school expenditure but the differential enrollment of boys versus girls. Using two rounds of nationally representative survey data, the paper examines the impact of a cashless, paperless and portable health insurance scheme called the Rashtriya Swasthya Bima Yojana (RSBY) launched in 2008 in India, on schooling decisions and gender differences in education. Employing difference-in-differences and triple differences approach, the paper finds that access to RSBY is beneficial for child education as school expenditure increases after the treatment. Additionally, RSBY is found to be relatively more advantageous to girls as it reduces the existing gender gap in school enrollment. Robustness checks and sensitivity analyses support the validity of the results.

JEL Codes: D13, I13, I18, I25, I26, I28

Keywords: Gender, education, health, health insurance, household decisions, RSBY, India

1 Introduction

Concerns about adequate healthcare and access to health insurance have witnessed profound growth over the past few decades amongst policymakers worldwide. The WHO states that 400

*I am deeply indebted to my adviser Daniel Millimet for his useful insights and guidance. I also thank Elira Kuka, Thomas Fomby, Santanu Roy, James Lake, Omer Ozak, Anil Swarup, Nishant Jain and fellow students of Southern Methodist University for helpful discussions and comments. All remaining errors are my own.

†Southern Methodist University, Department of Economics, 3300 Dyer Street, Suite 301, Umphrey Lee Center, Dallas, TX 75275-0496, United States. Email: mojha@smu.edu

million people in the world have no access to essential health services and 6 percent of people in developing countries are pushed further into extreme poverty due to health spending (WHO, 2015). Health shocks can be particularly devastating for the poor in developing countries owing to a lack of affordable insurance (Hamoudi et al., 1999, Wagstaff et al., 2009).[1] Absence of a formal pervasive public insurance system means large out-of-pocket expenditure is the main source of healthcare for them. As such, the burden of health shocks may be greater if its consequences are transferred to human capital of future generations in families unable to access formal insurance markets (Currie and Moretti, 2007; Bhalotra and Rawlings, 2011;Flores et al., 2008; Morduch, 1999; Sun and Yao, 2010).

Child human capital formation can potentially be affected through the following channels. First, if children are considered as substitutes for adult labour in a family with an ailing parent, they are compelled to be withdrawn from school and sent to work to smooth consumption (Fabre and Pallage, 2015).[2] Second, if the case is of an ailing child, they are withdrawn from school as their survival and health status assume more importance in such situations. Third, health shocks reduce a household's ability to afford the upfront cost of schooling. In the absence of safety nets coupled with poverty, households thus resort to financing healthcare expenditure through other costly measures like reduction in school expenditure or delaying their children's enrollments. However, it is not entirely obvious whether access to health insurance would have a positive or a negative impact on child education. On the one hand, the above mechanisms imply that insurance could protect children from being pushed into labour and reduce school dropouts in households affected by health shocks. On the other, better child health as a result of insurance could even mean more child labour for such families. The effect could thus be ambiguous and speaks to the importance of addressing it empirically.

Moreover, the impact of health insurance on education may not be gender-neutral. That, there exists a problem of gender gap in education in developing countries, is well known.[3] Researchers cite several reasons for this gap, like differential economic returns to education, parental preferences or biases, concerns over old-age support, and family's economic conditions, of which, health spending is a key determinant. Given this context, it is noteworthy to examine whether a health insurance system designed for the poor impacts schooling decisions and gender differences in education.

Health insurance has the potential to impact both the time opportunity cost of schooling as well

[1]Most private healthcare deliveries have low penetration due to lack of awareness and affordability. As a result, the government often fills this void in the market.

[2]They may even be asked to look after the sick parent reducing the time they can devote to school (Bratti and Mendola, 2014).

[3]Girls tend to receive less schooling than boys (Burgess and Zhuang, 2000; Schultz, 2002; Colclough et al., 2000; Alderman et al., 1996; Alderman and King, 1998).

as the monetary costs of schooling. Gender specific roles within the household result in higher time opportunity cost of investments in human capital for girls than boys.[4] Additionally, resource constraints exacerbate patterns of preferences within the household as income changes (Hill and King, 1995, Alderman and Gertler, 1997).[5] Basic education in developing countries is public but school attendance still requires out-of-pocket expenditures, sometimes large enough to keep children out of school. Although direct fee is unlikely to differ by gender, costs such as those of reaching school, learning materials, and uniforms may influence schooling decisions of girls more than boys. Access to a health insurance system for the poor could perhaps ease some resource constraints resulting in a reduction in the gender gap in enrollments.

This paper focuses on India's cashless, paperless and portable health insurance scheme started in 2008, called the Rashtriya Swasthya Bima Yojana (RSBY) to investigate these issues.[6] RSBY was implemented with the aim to protect the poor, across rural and urban areas, from financial liabilities and increase their access to quality healthcare. Given that 28 percent of India's population is below poverty line, health care expenditure is one of the most important reasons for indebtedness. Alarmingly, less than 15 percent of the 1.1 billion population are covered by health coverage. Moreover, over 78 percent of all medical expenditure in India is private financing most of which is out-of-pocket expenditure and is amongst the highest in the world (Swarup and Jain, 2011).[7] Initially targeted at below poverty line (BPL) households, RSBY has since expanded to cover other unorganized workers and marginalized sections who enroll into the scheme. The beneficiaries of the scheme are provided with a bio-metric smart card that can be used to receive health services from hospitals empanelled under the scheme without any out-of-pocket expenditure subject to certain conditions. RSBY therefore, assumes importance as a policy measure to not only decrease the vulnerability of credit-constrained households but to also potentially protect their children from adverse shocks.

While there exists literature on the impact of health insurance on health expenditure and health related outcomes in India, most papers focus on smaller insurance schemes concentrated in some states. Few recent empirical papers investigate the impact of RSBY on financial burden, health services and expenditures (Azam, 2018; Karan et al., 2017; Ravi and Bergkvist, 2015; Karan et al.,

[4] Girls invariably become the first victim of a health shock to the family without insurance (Garg and Morduch, 1998).

[5] Sons are valued more as they are considered labour assets and support during old age. Daughters however, usually leave the natal family post marriage (Sen and Sengupta, 1983, Bardhan, 1985; Rosenzweig and Schultz, 1982, Duraisamy, 1992, Garg and Morduch, 1998, Kingdon, 2005, Almond et al., 2010, Haddad et al., 1984).

[6] In recent times, many developing countries have subsidized health insurance for the rural and informal sector workers and their families (Wagstaff et al., 2009). China adopted a new health insurance system for the rural population called the New Cooperative Medical Scheme. On similar lines, Vietnam, Taiwan, Indonesia, and Philippines are also striving to achieve universal health coverage.

[7] External aid to the health sector accounts for a negligible 2 percent of the total health expenditure.

2014; Johnson and Krishnaswamy, 2012). However, thus far, no evidence exists for the spillover effects of health coverage, in general, and RSBY, in particular, on education. This is the first paper, to my knowledge, to investigate the role of a public health insurance scheme in India in determining school expenditure and enrollment decisions.

Using nationally representative longitudinal survey, my empirical analysis employs two different identification strategies. First, I estimate the effect of RSBY on both school expenditure and enrollment using a difference-in-differences strategy. Second, I employ a triple differences model which exploits the fact that rich households are significantly less likely to be affected by the program (due to the initial focus on BPL households). Using nationally representative *household* level data, I investigate the treatment impact of RSBY on household school expenditure. In addition, using nationally representative *individual* level data, I quantify a similar treatment effect of RSBY on school enrollment and the existing gender gap. I compare households in districts that are exposed to RSBY by the second wave of the survey, to those that were never exposed to the scheme in order to obtain the intent-to treat (ITT) impact of the programme.

The findings are interesting and ought to serve as a guide to future research and policy discussions. A key result is that access to health insurance is beneficial for child human capital formation, as school expenditure increases at the household level after the treatment. The estimates found imply an increase in the budget share of school expenditure of 0.5 to 0.7 percentage points. This effect is statistically and economically significant given that school expenditure accounted for 2.5 percent of the budget share for such households prior to RSBY. This amounts to an increase of 20 to 28 percent in its budget share after the treatment. Given that health insurance reduces uncertainty about occupational hazards, availability and access to RSBY mitigates costly choices a household may otherwise resort to, like reducing school expenditure. These results are robust to several alternative modeling choices.

Finding positive impacts on household school expenditure, the paper goes further to quantify the effects of RBSY on school enrollments of children within households. I find a clear reduction in the gender gap in school enrollment after implementation of RSBY. Absent the programme, school enrollment of boys is about 6 percentage points more than girls. I find that the probability of enrollment is 0.8 percentage points higher for boys and 2.7 percentage points higher for girls, after the programme went into effect. Thus, the gap in enrollment reduces by one-third. Triple differences approach confirms this result for relatively less well-off households. The results are robust to variations in income distribution.

Rest of the paper proceeds as follows. Section 2 presents a review on related literature. Section 3 provides the background and programme details of RSBY. Section 4 is divided into sub-sections:

4.1 describes the data, followed by the estimation and identification strategy for the analysis of school expenditure and school enrollment in subsections 4.2 and 4.3 respectively. Section 5 discusses the baseline results followed by robustness of the baseline models in Section 6. Section 7 presents sensitivity analysis of school expenditure and school enrollment. The paper ends with the conclusion in Section 8.

2 Literature

This paper contributes broadly to two bodies of literature. First, it contributes to the vast literature on the impact of public health insurance schemes. Effect of health coverage on uptake of treatment, out-of-pocket expenditures, in-patient and out-patient services in developing countries have been examined in Acharya et al., 2012; Wagstaff et al., 2009. Currie and Gruber, 1996; Chen and Jin, 2012; Liu and Zhao, 2014 study its impact on other health-related outcomes like health care disparity, health statuses of new born children, mothers and the elderly.

In the Indian context, the impact of health insurance, particularly RSBY, on various outcomes are found in Azam, 2018; Karan et al., 2017; Raza et al., 2016; Devadasan et al., 2013; Das and Leino, 2011; Palacios et al., 2011; Johnson and Krishnaswamy, 2012; Rajasekhar et al., 2011; Virk and Atun, 2015; Ravi and Bergkvist, 2015. Using panel data from the India Human Development Survey (IHDS), Azam (2018) utilizes a difference-in-differences with propensity score matching approach to estimate the average treatment impact (ATT) of RSBY on the beneficiary households. The paper uses both the household and the individual level data from the IHDS to investigate the impact on utilization of health services for short term and long term morbidity, total out-of-pocket expenditures, per capita in-patient and out-patient expenditures. Both Karan et al. (2017) and Johnson and Krishnaswamy (2012) use difference-in-differences with matching at household level to evaluate the ITT impact of RSBY using cross-section data from the national sample survey (NSS). Karan et al. (2017) find marginal decline in in-patient, out-patient out-of-pocket expenditures and budget share of out-of-pocket expenditure. Johnson and Krishnaswamy (2012) find that the scheme has led to a small decrease in out-patient and total medical expenditure of target households and some limited evidence of increased hospital utilization rates. On similar lines, Ravi and Bergkvist (2015) also use data from NSS and implement difference-in-differences across insurance districts versus uncovered districts to study the ITT impact of publicly provided health insurance schemes in India on the likelihood of impoverishment, catastrophic health expenditure, and the poverty gap index. Nandi et al. (2013) use district-wise official data on enrollment, and correlate those with district characteristics to find the determinants of participation in RSBY. Fewer studies have focused on the impact of health insurance on non health related outcomes. Among these papers, most have looked at the impact on household choices associated with health shocks (Kochar, 1995; Liu, 2016; Mohanan, 2013).

Second, the paper adds to the strand of literature on gender gaps in treatment of children in south Asia. According to some papers, boys are favored over girls in terms of intra-household allocation of resources and nutrients as found through indices like weight for age, mortality rates, and breast-feeding (Barcellos et al., 2014; Behrman, 1988; Bardhan, 1985; Sen and Sengupta, 1983; Rosenzweig and Schultz, 1982). Other papers suggest household income, parental education and supply side factors like quantity and quality of schools are explanations for low educational achievements and gender gaps in such countries (Behrman and Knowles, 1999; Duraisamy, 1992; Kambhampati and Pal, 2001; Pal, 2004; Dreze and Kingdon, 2001). More specifically, in the context of India, evidence of gender differences in child schooling exists for some states but very few studies are able to explain such differences (Pal, 2004; Glick et al., 2016).

3 Background on RSBY

Rashtriya Swasthya Bima Yojana (RSBY) or the national health insurance scheme was launched by the government of India as a cashless, paperless and portable health insurance scheme in 2008. The scheme was initially designed to target below poverty line population (BPL) both in rural and urban India but was later expanded to also cover unorganized workers such as construction workers, domestic help, street vendors, rickshaw pullers etc. RSBY aims to protect the poor from financial risk arising from out-of-pocket expenditures on hospitalizations and to improve the access to quality healthcare. Unlike most central government schemes, implementation of RSBY did not follow a top driven approach. The government marketed the scheme and rolled it out in districts based on factors such as need for the scheme, ease of implementation and acceptance from local governments. By October 2013, approximately 36 million families out of a target of approximately 65 million were enrolled in the scheme. As of 2013, the scheme was implemented in 512 districts out of 640 districts in 29 states across India (Government of India, 2013).

Beneficiaries of the scheme are entitled to hospitalization coverage of up to INR 30,000 (approximately \$460) for a family of five and transportation costs up to INR 1,000 (approximately \$16). The scheme is jointly funded by the central and state governments with 75% of premium from the center and 25% from the state.[8] State governments set up state agencies to prepare a list

[8] In case of Jammu & Kashmir and North-eastern States, 90% of premium is from the central government and 10% from the state.

of identified households.[9] Awareness campaigns are conducted through the Gram Panchayat and enrollment camps set up across districts.[10] Insurance companies, selected through a competitive bidding process by the government, are responsible for reaching out to the beneficiaries for enrollments. Once a hospital is empanelled, a nationally-unique hospital ID number is generated so that transactions can be tracked at each hospital.

Beneficiaries pay a small amount of INR 30 (approximately $5) as registration fee which is aggregated at the state level and is used to take care of the administrative cost of the scheme. Households that choose to enroll into the scheme receive a bio-metric card with a national unique ID. Upon receiving the card, the beneficiary can visit any empanelled hospitals across the country to get cashless treatment. Insurance companies are paid a fixed price per household enrolled and must settle all claims with the hospitals directly based on rates fixed by the central government. While all pre-existing diseases are covered, the scheme does not cover out-patient procedures. There is no age limit on the enrollment of beneficiaries.

4 Empirics

4.1 Data

I utilize two waves of the India Human Development Survey (IHDS), collected in 2004-05 and 2011-12 for the analysis.[11] IHDS is a nationally representative multi-topic survey of approximately 40,000 households across 1503 villages and 971 urban neighbourhoods of India. The surveys are collected from January to March.[12]

IHDS-II is mostly re-interviews of households interviewed for IHDS-I. I merge the two survey waves for my analysis both at *household* as well as *individual level.* The household sample is restricted to include households with children and where the age of head lies between 18 to 90 years. After dropping these observations, my sample consists of 29,381 households in the first survey wave and 25,226 in the second. The individual level sample is restricted to children in the age group 5 to 18

[9] These are referred to as the state 'nodal' agencies by the government.

[10] The date and location of the enrollment camp are publicized in advance. Some mobile enrollment stations are also established at local centers like public schools at each village at least once a year. These stations are equipped by the insurer with the hardware to collect bio-metric information and photographs of the members of the household covered.

[11] IHDS I refers to the time period 2004-05 and IHDS II to 2011-12.

[12] IHDS is collected by the National Council of Applied Research and Training (NCAER), New Delhi and University of Maryland. The waves are publicly available to be downloaded from the Inter-University Consortium for Political and Social Research (ICPSR). IHDS-I surveyed 41,554 households and IHDS-II 42,152 households.

years. The sample consists of 48,571 children in the first survey wave and 41,576 in the second. I consider the individual level data as repeated cross-section since it is difficult to track the same child over a period of 7 years between 2005 and 2012. Some children may have finished school while new children are enrolled. For consistency purposes, I also consider the household sample as a repeated cross-section.[13] I merge both the household and the individual samples separately with data on implementation of RSBY at district-level. Information about the roll-out of health insurance scheme is taken from the official ministry website.[14] The final sample consists of 393 districts across India. No districts were treated in the first wave and 53 districts were not treated as of the second wave. The summary statistics for my control and treatment districts in the two time periods are presented in Tables 1 and 2.

I consider budget share of school expenditure out of total monthly household expenditure as my outcome variable for the baseline analysis at *household level* (see Eqns. 1 and 2). Child specific school enrollment within each household is my outcome variable for the analysis at *individual level* (see Eqn. 3 and 4). Standard errors are clustered at district level in all estimations.[15]

The set of controls for household level analysis (refer to Eqns. 1 and 2) includes household size, age of the head of the household, age squared, educational characteristics of male and female members of the household, number of years a family has stayed in one place, indicators for caste (Brahmins, Scheduled Tribes, Scheduled Castes, and Other Backward Class), indicators for religion (Hindu, Muslim, Sikh, Buddhist, Jain, other religion), dummy for urban areas, whether head of the household can converse in English, gender dummy for the head of the household, number of married male and females in the households, dummies indicating number of years of marriage, whether the household has a bank account and a credit card. Control variables for the individual level analysis (refer to Eqn. 3 and 4) include household size, age of the child, age squared, mother and father's education characteristics, indicators for caste, religion dummies, dummy for urban areas, school facilities and scholarships offered. In addition to this I also include an indicator for the relatively poorer households, that takes value 1 if the household belongs to the bottom 70 percent of income distribution in my sample and 0 otherwise (refer to Eqn. 2 and 4 in 4.2 and 4.3).

Note that in all my models, I include household size as a regressor which is likely endogenous. Excluding household size as a control while analyzing school expenditures and gender differences in education implies that boys and girls live in families with similar characteristics, in terms of

[13]I redo my analysis at household level treating the household data as a true household level panel data for robustness purposes (refer to section 6.1.3).

[14]List of districts and phases of implementation can be found at http://www.rsby.gov.in/.

[15]This is true except when I estimate the treatment effect considering my sample as a panel data with household fixed effects. I cluster the standard errors at household level in this case.

both observables and unobservables. However, this assumption is likely to bias the estimates if families have a preference for sons and follow male-biased stopping rules of childbearing (Barcellos et al., 2014). If fertility decisions are driven by a desire to have a certain number of boys, then girls end up in larger families on average. To address this concern, I instrument household size by by gender of the first born child in the family under the assumption of no sex-selective abortion.[16] Although this assumption is not without criticism, in such cases, gender of the first child is likely to be a good predictor of the number of children in the household or family size and excludable from the second stage (Barcellos et al., 2014,Clark, 2000,Clarke, 2017).

4.2 School expenditure - Estimation and identification

I use a difference-in-differences (DID) strategy to compare households in districts that are exposed to health insurance by the second wave of the survey to those that were never exposed to the scheme. All households in 2004-05 and some households in 2011-12 that are never exposed to RSBY form my control group and the households in districts exposed to RSBY in 2011-12 form my treatment group. A simple comparison of households from districts that received the scheme to those that did not would likely lead to biased estimates. I include district fixed effects to address the concern of any time invariant district level characteristics that may be correlated with the treatment. Time fixed effects control for the time-varying characteristics that impact all districts equally. Identification relies on changes in household school expenditure at the district level after the phase-wise implementation of RSBY in 2008. I am unable to identify which households directly participated in the programme. Thus, I use all the households in a treated district and estimate the effect of access to the programme. This is the intent-to-treat (ITT) effect of RSBY on school expenditure.

I use the following DID specification to compare the households in districts over the two time periods, 2004-05 and 2011-12, before and after RSBY was rolled out:

$$y_{hdt} = \beta_0 + \beta_1 T_t + \beta_{DD} RSBY_{dt} + \gamma X_{hdt} + \mu_d + \epsilon_{hdt} \qquad (1)$$

where y_{hdt} is the budget share of school expenditure in household h in district d at time t. T_t takes the value 1 for 2011-12 and 0 for 2004-05. $RSBY_{dt}$ a treatment indicator which takes the value 1 if district d is exposed to RSBY in time t and and 0 otherwise.[17] X_{hdt} is the set of household level

[16] Ban on sex-selective abortion was enacted in India in 1971 and later amended in 2004 making prenatal sex-screening and sex-selective abortion punishable by law (United Nations, 2017)

[17] Note that $RSBY_{dt}$ varies with both district and time and is equivalent to the usual $treat \times post$ that one finds in difference-in-differences analyses.

controls and μ_d depicts district fixed effects. The disturbance term ϵ_{hdt} summarizes the influence of all other unobserved variables that vary across households, districts, and time. The baseline Eqn. 1 is estimated using an Instrumental Variable approach (IV).[18] The parameter of interest is β_{DD} which provides the differential impact of RSBY on household's expenditure on school after its introduction. β_1 identifies the effect of any systematic changes that affected households in all districts between 2004-05 and 2011-12.

A primary concern with the identification strategy in a DID approach is that the districts may be trending differently prior to RSBY. Ideally, two rounds of survey waves prior to the scheme would aid in analyzing the pre-trends. However, this survey was conducted for the first time in 2004-05, which restricts my analysis of pre-trends in expenditure. To alleviate such concerns, I estimate a triple differences model where I refine the definition of my control and treatment groups. I include the indicator variable $LowInc_h$ for poorer households as described in 4.1. Households in the top 30 percent are now the controls for such differential trends in the districts. The assumption here is that the richer households are perhaps less affected by RSBY and the two groups have similar trends in expenditure. This is reasonable since richer households are less likely to be resource constrained and in a position to insure themselves against unexpected shocks or have access to private health insurance.[19]

As such, the triple differences (DDD) estimator is more convincing as it looks at changes among poorer households in treated versus the control districts and nets out any differential change in wealthy households across treated versus control districts. The main identification assumption in such triple differences model is no longer that changes in treatment households should be uncorrelated with district level trends, but that these changes should be uncorrelated with district level trends that affect the rich and the poor differently. The assumption in this model is indeed weaker. This methodology helps take care of two potential confounding elements that are of concern in a DID model. One, the changes in school expenditure of the poorer households in the treated districts is not a result of changes in school expenditure of such households across all districts, nor is it a result of changes in school expenditure of all households in the treatment districts (possibly due to other unobservables that affects all households).

[18] Gender of the first born child in the family is used as an instrument for household size.

[19] It must be noted that the initial target population intended by the scheme was the bottom 30 percent of income distribution. However, the scheme was later extended to several unorganized workers over the years (Government of India (2013)). At the outset, it is necessary to caution that the top 30 percent may not form a clean control. I test the robustness of my triple difference results by altering the income distribution categories for my control and treatment groups. These are discussed in the later sections.

The second specification I estimate is the following triple differences model:

$$y_{hdt} = \beta_0 + \beta_1 T_t + \beta_2 LowInc_h + \beta_3 RSBY_{dt} + \beta_4 T_t * LowInc_h + \beta_{DDD} RSBY_{dt} * LowInc_h + \mu_d * LowInc_h + \gamma X_{hdt} + \mu_d + \epsilon_{hdt} \quad (2)$$

where $RSBY_{dt}$ is the treatment dummy that varies with district and time. The new coefficient of interest is β_{DDD} which is the difference-in-difference-in-differences estimator. β_{DDD} captures the variation in school expenditure in poorer households (relative to the rich) in treated districts (relative to control districts) after implementation of RSBY. Similar to the DID model, the baseline triple differences in Eqn. 2 is also estimated using an IV approach. Other set of controls are same as the baseline model. District fixed effects, time fixed effects, time by income fixed effects and district by income fixed effects are included. Standard errors are clustered at district by household income level.[20]

4.3 School enrollment - estimation and identification

Ideally, investigating within-household expenditure patterns on boys versus girls would help quantify the exact gender differences in parental investment. However, studies that have attempted to examine gender bias in schooling through household expenditure data have met with little success. Expenditure on individual members of a household is typically not observed in survey data which makes it impossible to directly observe gender biases in allocation of expenditure. Most papers therefore, resort to indirectly detecting differential treatment within households by examining changes in household expenditure with changes in gender composition. Reliability of this methodology however, has been called into question because it generally fails to detect a gender bias (Deaton, 1997). Even in countries with known gender bias, researchers thus far find mixed evidence of significant effects of the child's gender on the composition of household spending (Bhalotra and Attfield, 1998). Similar lack of convincing expenditure data at the child level makes it impossible for me to quantify the treatment impact on gender differences in parental investments in educational expenditure. Instead, I use data at individual level on school enrollments to get at the treatment effect on gender differences in boys' and girls' enrollments within households.

[20] For comparison purposes, I also estimate both the DID and DDD models for the numerator and denominator of the budget share separately, that is, logarithm of school expenditure in levels and logarithm of total consumption expenditure in levels for the household. This is discussed in the robustness section 6.1.1.

I estimate the following linear probability model (LPM) to estimate the treatment effect

$$y_{ihdt} = \alpha_0 + \alpha_1 T_t + \alpha_{DD,i} RSBY_{dt} + \gamma X_{ihdt} + \mu_d + \epsilon_{ihdt}$$

$$where \;\; \alpha_{DD,i} = \alpha_2 + \alpha_3 Boy_i$$

$$\implies y_{ihdt} = \alpha_0 + \alpha_1 T_t + \alpha_2 RSBY_{dt} + \alpha_3 RSBY_{dt} * Boy_i + \gamma X_{ihdt} + \mu_d + \epsilon_{ihdt} \quad (3)$$

where y_{ihdt} is an indicator variable which takes value 1 if the child i in household h in district d is enrolled in school in time period t. Boy_i takes value 1 if the child is a boy and 0 if a girl. District and time fixed effects are included in the model and standard errors are clustered at district level. α_1 identifies the effect of any systematic changes that affect the child between the two time periods. α_2 depicts school enrollment of a girl as a result of the treatment. $\alpha_2 + \alpha_3$ identifies the school enrollment of a boy post the treatment. The coefficient of interest is α_3 which gives the change in the gender gap in school enrollment due to RSBY. I also control for the gender dummy of the child, the coefficient of which identifies the school enrollment of boys versus girls absent the treatment. All other relevant controls are included as described in the section 4.1.

Similar to the school expenditure triple differences analysis, I also estimate an equivalent model for school enrollment. Incorporating the new treatment and control groups, the specification looks as follows:

$$\begin{aligned} y_{ihdt} = {} & \alpha_0 + \alpha_1 T_t + \alpha_2 LowInc_h + \alpha_3 RSBY_{dt} + \alpha_4 T_t * LowInc_h + \alpha_{DDD,i} RSBY_{dt} * LowInc_h \\ & + \mu_d * LowInc_h + \gamma X_{ihdt} + \mu_d + \epsilon_{ihdt} \end{aligned}$$

$$where \;\; \alpha_{DDD,i} = \alpha_5 + \alpha_6 Boy_i$$

$$\begin{aligned} \implies y_{ihdt} = {} & \alpha_0 + \alpha_1 T_t + \alpha_2 LowInc_h + \alpha_3 RSBY_{dt} + \alpha_4 T_t * LowInc_h + \alpha_5 RSBY_{dt} * LowInc_h \\ & + \alpha_6 RSBY_{dt} * LowInc_h * Boy_i + \mu_d * LowInc_h + \gamma X_{ihdt} + \mu_d + \epsilon_{ihdt} \quad (4) \end{aligned}$$

where α_5 depicts the effect of RSBY on enrollment of girls and $\alpha_5 + \alpha_6$ depicts the effect of RSBY on enrollment of boys. Change in the gender gap in school enrollments as a result of RSBY for poorer households in the treated districts is thus given by α_6. It captures the variation in boys' and girls' school enrollments within such households in the treatment districts, nets out the change in the average enrollments in the control districts and then nets out the change in the average enrollments in richer households in the treatment district. As before, the model includes all controls, all relevant double interaction terms as well as district and time fixed effects.

5 Results

5.1 School expenditure as a budget share

I present the baseline school expenditure results in Table 3. Panel A presents the results for the DID specification 1. Column (1) shows that RSBY increases the budget share on school expenditure by 0.5 percentage points and the effect is statistically significant at p<0.01 significance level. Access to health insurance has positive spillover effect on school expenditure decisions of households. Panel B presents the results for the triple differences estimation of Eqn. 2. From column (4), notice that the triple differences analysis gives a treatment effect of the order of 0.7 percentage points on the budget share of school expenditure for the poor households relative to the rich in treatment district relative to control. Columns (2), (3), (5) and (6) present the impact of RSBY on the logarithm of school expenditure in levels and logarithm of total consumption expenditure in levels for DID and DDD models.[21]

Summary statistics in table 1 shows that the average share of school expenditure out of total expenditure for such households in 2004-05 is about 2.5 percent. Both the DID and DDD effects are therefore economically significant and imply that the budget share of school expenditure increases by 20 to 28 percent after RSBY. To the extent that access to public health insurance helps reduce household's financial burden, RSBY benefits child human capital formation through an increase in expenditure on school. As such, RSBY perhaps helps eliminate costly smoothing mechanisms that households may resort to, in absence of such an insurance coverage, like cutting down on school expenditure or delaying their children's enrollments.[22]

Note that, several diagnostic tests have been performed to assess the efficiency and reliability of the instruments. The endogeneity test reports test statistics that are robust to various violations of conditional homoskedasticity. I reject exogeneity of household size.[23] As far as underidentification is concerned, I report chi-squared p-values for the test where rejection of the null implies full rank and identification (Baum et al., 2007). This test tells us whether the excluded instrument is correlated with the endogenous regressor. The p-value based on Kleibergen-Paap rk LM statistic allows

[21] These results are discussed in detail in section 6.1.1.

[22] Selling assets, exhausting savings, non-institutional borrowings and reducing consumption below critical levels are other examples of such costly measures (Morduch, 1999,Sauerborn et al., 1996; Edmonds, 2006).

[23] Under conditional homoskedasticity, this endogeneity test is numerically equal to a Hausman test statistic.

me to clearly reject the null that the instrument is uncorrelated with the endogenous regressor and that the model is underidentified. From the weak identification test, rejection of the null represents absence of weak-instrument problem. Since the specification has clustered standard errors at district level, the reported test statistic is based on the Kleibergen–Paap rk statistic which indicates absence of weak instrument problem, given that it is above 10 in the baseline specification of DID (column (1)).[24]

5.2 School enrollment

Given that RSBY has an impact on budget share of school expenditure at the household level, it is noteworthy to examine its impact on gender gap in school enrollments within households. I present the results for the baseline school enrollment analysis in Table 4. Panel A provides the DID results estimated using a linear probability model for specification 3. Column (1) presents the impact on enrollments without a gender differential whereas column (2) presents the impact when I introduce a gender differential. In this case, notice that absent the health coverage, a gender gap in school enrollment exists. More boys are enrolled in school. In fact, enrollment of boys is about 6 percentage points higher than that of girls. Average enrollment is 78.4 percent for boys and 72.4 percent for girls prior to the treatment. Difference in parental expected future returns from their children's schooling or parental preferences could be possible explanations for this, as found in extant literature. If parents expect higher returns from boys than girls, it limits the amount of equality a household can afford. Column (2) shows that I find the treatment to have a larger impact on girls. The probability of enrollment is 2.7 percentage points higher for girls after implementation of RSBY as compared to 0.8 percentage points higher for boys. The reduction in the gender gap as a result of the treatment is by 1.9 percentage points and is statistically significant at $p<0.01$ significance level. The triple differences results for specification 4 are presented in panel B. Column (4) shows a reduction (albeit smaller in comparison to DID) in the gender gap in enrollment by 0.9 percentage points and is statistically significant at $p<0.05$ significance level. This suggests that benefits of the health insurance scheme accrues more to girls insofar as school enrollment is concerned.

Gender specific roles in domestic chores and differential time opportunity cost of boys' and girls' schooling explains these results to some extent. As suggested, differential patterns of preferences within the household are exacerbated with changes in household income (Alderman and Gertler, 1997). Given that girls spend less time in school and more hours working to substitute for mothers'

[24]The instrument becomes slightly weaker in the baseline of triple differences model owing to perhaps more number of controls and lower correlation.

domestic duties, the greater impact on girls could perhaps be a result of RSBY reducing the degree of impact of a shock to mother's health on daughters.[25] One could perhaps also say that larger treatment effect on enrollment of girls is because the demand for girls' human capital is more income and price elastic than demand for boys'. Moreover, although basic education in India is tuition-free, school attendance still entails cost of reaching school, learning materials, uniforms that are large enough out-of-pocket expenditures to keeps more girls out of school. Access to a cashless health insurance system perhaps eases some resource constraints in the households leading to a reduction in the gender gap in enrollments post the treatment.

6 Robustness Checks

There may be other potential concerns related with my baseline estimations. This section discusses the additional analyses I conduct to explore the robustness of my results to different modeling choices for both school expenditure as well as school enrollments. I start with a discussion of school expenditure models and then proceed to school enrollments.

6.1 School expenditure - other estimation issues

Taking the budget share of household school expenditure as my outcome variable would ideally require me to estimate a fractional response model.[26] However, given that I am controlling for a large number of districts, a fractional response model with fixed effects becomes infeasible. I therefore, compare the baseline IV results with those from two main alternative estimation approaches.

6.1.1 School expenditure in levels

My treatment effect could possibly be understated when I consider budget share of household school expenditure as the dependent variable. A direct positive income effect of RSBY could perhaps be translated to an increase in total household consumption expenditure itself given that health insurance relieves household's resource constraints. If total consumption expenditure of households rises, this would mean a lower effect on the budget share of school expenditure. Therefore, I first

[25] With women receiving less healthcare, a shock to the mother's health would have a larger impact on the girls required to take up on mother's chores (Alam, 2015, Hazarika and Sarangi, 2008, Katz, 1995, Skoufias, 1993)

[26] The budget share is a fraction and is bounded between 0 and 1.

estimate a model where the outcome variable is the logarithm of household's school expenditure per month in *levels* excluding total consumption expenditure from the specification. I also estimate the treatment effect on logarithm of total consumption expenditure in *levels*. This helps me tease out the treatment effect on both household school expenditure and total consumption expenditure separately.

Note that in IHDS survey, some households report zero expenditure on goods. My dependent variable is in logarithms which implies that value of the corresponding outcome variable will be undefined if I include such households. One way to avoid this problem is simply to drop these households and run regressions based on the trimmed sample. However, this may result in sample selection bias. Rather, a more sophisticated way to circumvent this problem and include these households is to apply the *inverse hyperbolic sine transformation* of consumption expenditures (Burbidge et al., 1988). The inverse hyperbolic sine transformation requires transformation of the variable in question, say, z as $log(z^2 + \sqrt{z^2+1})$ which unlike $log\, z$, is defined even for $z = 0$.[27] As such, in this paper I use the inverse hyperbolic sine transformation to deal with households reporting zero consumption expenditure.

The results for these is presented in Table 3. Columns (2) and (5) provide the DID and DDD effects on log of school expenditure in levels. I find that RSBY increases school expenditure by 30.2 to 42.2 percent approximately. The effect is found to be greater in the DDD model for the poorer households in treated districts. Column (3) provides the DID effect on log of total consumption expenditure. An increase by 7.7 percent is seen from column (3). I find a positive impact on log of total consumption expenditure in the DDD model as well but the effect is not statistically significant (column(6)).

Second, I estimate the *levels* model while controlling for total consumption expenditure as a regressor. This takes care of any income effect of the scheme as it holds the budget constraint constant for the household. However, there may be a possible endogeneity concern for total consumption expenditure here. I instrument total monthly household consumption expenditure by assets possessed by the household at the time of the survey to circumvent this problem. This serves as valid instrument because assets held at the time of the survey do not directly impact the monthly expenditure on school but are a good predictor of total household income or consumption. Monthly expenditures on commodities are usually out of current earned income rather than out of assets or wealth.[28]

[27]According to Burbidge et al. (1988), except for very small values of z, the transformation is approximately equal to $log(2z_i)$ or $log(2) + log(z_i)$, and so it can be interpreted in exactly the same way as a standard logarithmic dependent variable.

[28]Although, land could affect school expenditure to some extent since land requires work and missing work would factor into opportunity cost of expenditure related to school.

Panel A and B, Table 5 present these results. Columns (1) and (3) repeat my baseline results as in table 3. Columns (2) and (4) present the results where I include total consumption expenditure and instrument it with total household assets. In this specification, I have two endogenous regressors and two instruments. From column (2), the treatment effect shows an 8 percent increase in the level of school expenditure and is statistically significant while holding the budget constraint of the household constant. The triple differences model also shows a higher and statistically significant impact on the level of school expenditure of almost 18.7 percent for the poor households in the treated districts (see column (4)). Here, the total treatment impact from the triple differences model is 8.2 percent which is approximately equivalent to the difference-in-differences result. As before, diagnostic tests have been performed to assess the efficiency and reliability of the instruments. The instruments fair broadly well on these specification tests.

6.1.2 School expenditure - fractional logit estimation

Here, I return to budget share as my outcome but estimate a fractional response model with correlated random effects to account for district level characteristics since a fixed effects fractional response model is not feasible. I estimate specification 1 via a fractional logit model with correlated random effects (CRE). The advantage of using CRE fractional logit is that it places some structure on the nature of correlation between the unobserved effects and the covariates (Lake and Millimet, 2016). Formally, the structural model in the CRE fractional logit is given by

$$\mathrm{E}(y_{hdt}|\, X_{hdt}, \mu_d) = \Phi(X_{hdt}\beta + \mu_d) \tag{5}$$

where X_{hdt} includes the full set of covariates in specifications 1 and 2 and Φ is the standard normal cumulative density function. The Mundlak (1978) version of the CRE probit model further assumes

$$\mu_d|X_{hdt} \sim \mathbf{N}(\delta_0 + \bar{X}_h\delta_1, \sigma^2_\mu) \tag{6}$$

where $\bar{X}_h$ is the average of X_{hdt} for each district and σ^2_μ is the variance of μ_d.Under 5 and 6, we get

$$\begin{aligned}\mathrm{E}(y_{hdt}|\, X_{hdt}, \mu_d) &= \Phi[(\delta_0 + X_{hdt}\beta + \bar{X}_h\delta_1).(1+\sigma^2_\mu)^{-1/2}] \\ &= \Phi[\delta^\mu_0 + X_{hdt}\beta^\mu + \bar{X}_h\delta^\mu_1]\end{aligned} \tag{7}$$

To capture the district fixed effects in 7, means of all controls at district level across time are included as additional controls in the DID model. Standard errors are clustered at the district level and time fixed effects are included. I include the means of all controls at district by household income level as the correlated random effects for my triple differences model. Here, the standard errors are clustered at district by household income level.

Following Wooldridge et al., 2011; Wooldridge, 2015; Baum et al., 2013; Papke and Wooldridge, 2008, I use a two step *control function approach* to deal with the continuous endogenous regressor, household size included in my model. In the control function approach, I first estimate household size as a function of my instrument, which is, gender of the first child in the household. This gives me residuals similar to the first stage of a 2SLS approach. I then use the residuals from this model as an additional regressor in the main model which is estimated as a CRE-fractional logit model.

I present the results in Table 6. Panel A provides the DID results and panel B, the triple differences results. Columns (1) and (3) repeat my baseline results as in table 3. Columns (2) and (4) presents the results using IV results for the CRE-fractional logit model. Since column (3) is the CRE fractional logit specification, I cannot interpret the coefficients and thus calculate the marginal effect of the treatment. I find a small positive marginal effect of RSBY but it is not statistically different from zero.

The CRE fractional logit specification of triple differences model in column (4) shows a small but statistically significant difference in the marginal effects of RSBY for the poor and the rich households in the treated districts after RSBY. There is no effect on the rich households. The magnitude of this difference is of 0.2 percentage point which implies a difference of 8 percent in the budget shares for the poor and the rich in treated districts.

6.1.3 School expenditure - panel analysis

As an additional robustness check of my baseline school expenditure model, I estimate the treatment effect by considering the data as a panel since IHDS II are re-interviews of most of IHDS-I households. The results are overall robust to this change. Table 7 presents the results. Panel A provides the difference-in-differences results and panel B, the triple differences. Columns (1) and (3) repeat the baseline DID and DDD results as in table 3. Column (2) shows the effect of RSBY using IV approach with household fixed effects for the panel data. The standard errors are clustered at household level. RSBY increases budget share of school expenditure by 0.3 percentage points as suggested by the DID model. This implies a 12 percent increase in the budget share of school expenditure given that the mean budget share was 2.5 percent from Table 1. From Column (4), the triple differences estimator shows that RSBY leads to an increase of 0.4 percentage points in the budget share of school expenditure for the poorer households in the treated districts. This is equivalent to a 16 percent increase in the budget share of school spending for the poorer households.

6.2 School enrollment - other estimation issues

6.2.1 School enrollment - probit with correlated random effects

A first potential concern with my school enrollment analysis is that the dependent variable is a binary outcome and should ideally be estimated as a non-linear model such as a probit or a logit. Linear probability models are likely to give a biased and inconsistent estimate (Horrace and Oaxaca, 2006). Probit or logit models however use a proper functional form where the probability depends on x through the index $x\beta$

$$\Pr(y_i = 1|x_i) = \mathrm{F}(x_i\beta)$$

where the functional form $\mathrm{F}(\cdot)$ maps into a response probability $\mathrm{F} : \mathbb{R} \longrightarrow [0, 1]$ for which we consider CDFs as they map numbers from the entire real number line on to the unit interval. Given that the difference between a probit or a logit is small in practice, I use a probit model. However, as before, I have a fixed effects baseline model where I control for a large number of districts making a simple probit estimation infeasible. Thus, I compare estimates from the baseline fixed effects linear probability model with two alternative estimation approaches to analyze the robustness of my results. First, a linear probability model with correlated random effects. Second, an IV probit model with correlated random effects for which, a variant of 7 would look like

$$\Pr(y_{ihdt} = 1|\, X_{ihdt}, \mu_d) = \Phi[(\delta_0 + X_{ihdt}\beta + \bar{X}_i\delta_1).(1 + \sigma_\mu^2)^{-1/2}] = \Phi[\delta_0^\mu + X_{ihdt}\beta^\mu + \bar{X}_i\delta_1^\mu] \qquad (8)$$

Again, to capture the district fixed effects, means of all controls at district level across time are included as additional controls in the estimation. All standard errors are clustered at the district level and time fixed effects are included.

The results are presented in Table 8. Panel A presents the difference-in-differences results and panel B, the triple differences results. One can compare the baseline results presented in columns (1) and (4) with those from a linear probability model with correlated random effects presented in columns (2) and (5) as well as CRE-IV probit model presented in columns (3) and (6). Notice that from both the linear probability models with fixed effects and correlated random effects, the DID approach show that, absent the treatment, the probability of enrollment of a boy is approximately 6 percentage points higher than a girl. After the treatment, I find a larger effect on probability of girls' enrollment. A reduction in the gender gap in enrollment of 1.8 percentage points is seen from column (2). Since column (3) presents the probit model results, I cannot simply interpret the coefficients. Looking at the marginal effects of RSBY, I find consistent results. Notice that marginal effect of the treatment on the probability that enrollment of a girl is statistically significantly higher than that of a boy. The DID estimation shows that reduction in gender gap as a result of access to health insurance is of 3.2 percentage points and statistically significant.

The triple differences results confirm a similar story. Results from both a LPM with fixed effects and LPM with CRE are quantitatively similar. The impact of RSBY on the probability of enrollment of girls is higher. The reduction in enrollment gender gap of 0.9 percentage points is seen in both specifications. I also find a reduction in the gender gap in enrollment of 1.4 percentage points from the CRE-IV probit as seen by the marginal effects of RSBY on a boy and a girl in column (6), however, the effects are not precisely estimated.

6.2.2 School enrollment - instrumental variable approach

The second concern is related to the instrument used for household size in my school enrollment analysis. I compare how my results change from the baseline LPM model where I instrument household size with gender of the first born in the family, with two alternative specifications. First, I include household size in the specification but do not use an instrument for it. Second, I exclude household size from the specification.

In Table 9, I present the baseline LPM results in columns (1) and (4) and compare with LPM specifications where household size is included but not instrumented for (columns (2) and (5)) as well specifications where I omit household size as a regressor (columns (3) and (6)). Panel A presents the difference-in-differences results and panel B presents the triple differences results. Strikingly, the DID results from all three columns (1), (2) and (3) are qualitatively and quantitatively similar as well as statistically significant. I find a reduction in the gender gap in enrollment as a result of RSBY, of 1.8 to 1.9 percentage points, in all three estimation choices. The triple differences results also confirm a statistically significant reduction in gender gap in enrollment of 0.8 to 0.9 percentage points as a result of RSBY from all three specifications (columns (4), (5) and (6)). These alternative specification choices thus support validity of my baseline results.

7 Sensitivity Analysis

This section first discusses the sensitivity of my baseline results to variations in income distribution introduced in my models in 7.1. Second, I explore the heterogeneous effects of the treatment by intensity in section 7.2. Third, I exploit the variation in take-up of RSBY by district to estimate the heterogeneous treatment effect in 7.3. Lastly, I explore whether my baseline effects are different across sub samples varied by age groups for enrollment; by areas and by castes for both expenditure and enrollment in 7.4.

7.1 Variation in income distribution

I introduce variation in the income distribution categories used to define treatment and control groups in the triple differences model. To maintain symmetry with my baseline triple differences, I first restrict the sample to households in the top 30 percent and bottom 30 percent of income distribution. For this, I redefine $LowInc_h$ in Eqn 2 such that the top 30 percent households form controls for my new treatment group, which is, the bottom 30 percent. Observations in the middle 40 percent are dropped. Second, I drop observations from the bottom 30 percent and re-define $LowInc_h$ such that the top 30 percent are now controls for households in the middle 40 percent of the sample. Since RSBY was expanded to cover other unorganized and domestic workers, the expectation for this second variation in treatment group is that the effect is perhaps positive, but smaller.

I present these results in table 10. Panel I provides the results for school expenditure analysis. Panel A repeats the baseline DID results. Panel B presents the results for the two variations in my triple differences model. Column (2) shows that RSBY has a treatment effect of 0.5 percentage points increase in the budget share of school expenditure for the households that belong to the bottom 30 percent in the treated districts. This was the initial target group of the scheme. Average budget share of school expenditure for this target group in 2004-05 is approximately 1.8 percent. A treatment effect showing 0.5 percentage point increase thus implies approximately 27 percent rise in their budget share of school spending. Contrary to the expectation, the triple differences estimator in column (3) shows a zero effect on the households in the middle 40 percent of the sample. This could perhaps be a result of difference in the take-up of the program as the sample for this specification changes.

An equivalent school enrollment analysis is presented in panel II. RSBY leads to small reduction in the gender gap in enrollment of 0.2 percentage points for the bottom 30 percent in the treated districts. However, I do not find any reduction in the gender gap for the middle 40 percent households.

7.2 Variation in treatment intensity

Here, I exploit the variation in treatment intensity to estimate the heterogeneous effect of RSBY on household school expenditure as well as school enrollment. I define a ordered variable *Intensity* that takes values 1, 2, and 3 depending on the duration a district has been covered by RSBY until the second wave of the survey.[29] One may expect the effect of RSBY to vary with time since implementation. To explore this, I use the following difference-in-difference and triple differences models for school expenditure:

$$y_{hdt} = \beta_0 + \beta_1 T_t + \beta_2 RSBY_{dt} + \beta_3 Intensity_d + \beta_4 RSBY_{dt} * Intensity_d + \gamma X_{hdt} + \mu_d + \epsilon_{hdt} \quad (9)$$

$$\begin{aligned} y_{hdt} = \beta_0 + \beta_1 T_t + \beta_2 LowInc_h + \beta_3 Intensity_d + \beta_4 RSBY_{dt} + \beta_5 T_t * LowInc_h \\ + \beta_6 RSBY_{dt} * LowInc_h + \beta_7 RSBY_{dt} * LowInc_h * Intensity_d \\ + \mu_d * LowInc_h + \gamma X_{hdt} + \mu_d + \epsilon_{hdt} \end{aligned} \quad (10)$$

The parameter of interest varies with time t and district d, where the total impact of RSBY is given by $\beta_2 + \beta_4 Intensity_d$ in Eqn. 9. The heterogeneous effect of RSBY is captured by β_4. Similarly the parameter that captures the heterogeneous effect of RSBY in the triple differences Eqn. 10 is β_7 and the total effect of RSBY for the poor households is given by $\beta_4 + \beta_6 + \beta_7 Intensity_d$.

Similarly, I estimate the following DID and DDD models for school enrollment:

$$\begin{aligned} y_{ihdt} = \alpha_0 + \alpha_1 T_t + \alpha_2 Boy_{ihdt} + \alpha_3 RSBY_{dt} + \alpha_4 Intensity_d + \alpha_5 RSBY_{dt} * Boy_{ihdt} \\ + \alpha_6 RSBY_{dt} * Intensity_d + \alpha_7 RSBY_{dt} * Intensity_d * Boy_{ihdt} + \gamma X_{hdt} + \mu_d + \varepsilon_{ihdt} \end{aligned} \quad (11)$$

$$\begin{aligned} y_{ihdt} = \alpha_0 + \alpha_1 T_t + \alpha_2 Boy_{ihdt} + \alpha_3 LowInc_h + \alpha_4 RSBY_{dt} + \alpha_5 Intensity_d + \alpha_6 RSBY_{dt} * LowInc_h \\ + \alpha_7 RSBY_{dt} * LowInc_h * Boy_{ihdt} + \alpha_8 RSBY_{dt} * LowInc_h * Intensity_d \\ + \alpha_9 RSBY_{dt} * LowInc_h * Intensity_d * Boy_{ihdt} + \alpha_{10} T_t * LowInc_h \\ + \mu_d * LowInc_h + \gamma X_{ihdt} + \mu_d + \varepsilon_{ihdt} \end{aligned} \quad (12)$$

In Eqn. 11, $\alpha_3 + \alpha_6 Intensity_d$ provides the heterogeneous effect of the treatment on enrollment of girls by intensity of treatment duration whereas $(\alpha_3 + \alpha_5) + (\alpha_6 + \alpha_7) Intensity_d$ captures the

[29] By the second wave of the survey, the scheme was active for three years.

heterogeneous effect on enrollment of boys by intensity. The change in the gender gap due to RSBY is captured by $\alpha_5 + \alpha_7 Intensity_d$. Similarly, the change in the gender gap in enrollment due to RSBY for the poorer households in Eqn. 12 is given by $\alpha_7 + \alpha_9 Intensity_d$.

Results for the heterogeneous effects of RSBY by intensity of treatment are provided in Table 11. Panel A and B provide the DID and DDD results respectively. Panel I and II provide the results for school expenditure and school enrollment respectively.

Column (1) in panel I shows a treatment impact of 0.5 percentage point increase in budget share of school expenditure for households in districts that are exposed to the scheme for one year. This is equal to my baseline result found in column (1) Table 3 which implies that on average the treatment effect is equivalent to having been exposed to RSBY for a period of one year. The treatment effect is of the order of 1 percentage points and 1.5 percentage points for households in districts that have been exposed to the scheme for two and three years respectively. The DDD results do not show a statistically significant effect in this model but the effects point to a similar story.

Column (1) in panel II shows that the reduction in gender gap for the individuals in districts that have been exposed to RSBY for one year is by 3.3 percentage points. Consequently, this effect is of the order of 5.5 and 7.7 percentage points in districts that have had RSBY for two and three years respectively by 2011-12. The DDD analysis confirms this pattern. Column (2) in panel II shows a reduction in the gender gap in enrollments for boys and girls in poor households in districts exposed to RSBY for one year is by 2.6 percentage points. This result is a statistically significant and the gender gap consequently increases for such households that have had longer access to the scheme.

7.3 Variation in programme take-up by district

Third, I exploit district variation in the take-up of the programme to estimate the effect of RSBY on household school expenditure. Administrative reports suggests that health insurance take-up reached approximately 50 percent by 2013. Considering this, one would expect the treatment effect to be double if full take-up could be achieved. To explore this, I use the following difference-in-difference model

$$y_{hdt} = \beta_0 + \beta_1 T_t + \beta_2 RSBY_{dt} + \beta_3 DistrictTakeup_d + \beta_4 RSBY_{dt} * DistrictTakeup_d + \gamma X_{hdt} + \mu_d + \epsilon_{hdt}$$

The coefficient of interest is $\beta_2 + \beta_4$. Data for district-wise enrollment into the scheme is taken from the official RSBY website. RSBY enrollment data is available for districts from 15 states out of 29 is either because some districts have not been exposed to the scheme or simply because of unavailability of data.

Table 12 presents the results. Panel A, column (1) shows the simple difference-in-differences treatment results without differential take-up for the districts with available data. I find an RSBY leads to 0.7 percentage point increase in the household's budget share of school expenditure for the districts enrolled into the scheme. This is equivalent to a 28 percent increase in their budget share of school spending given that the school expenditure comprised approximately 2.6 percent of the total household expenditure for such households before RSBY. Column (2) in panel A provides the differential treatment of RSBY effect by take-up. If the treatment effect is extrapolated to a 100 percent take-up, I find household's budget share of school expenditure increases by 0.9 percentage points which is equivalent to almost 35 percent increase in the budget share of school spending for such the treated households. This is an economically large effect and is of interest since my treatment is the availability of the scheme and not household participation. However, a word of caution is warranted here that this is only suggestive evidence of the treatment effect since it is based on incomplete data and enrollment into the scheme is endogenous. In addition, my instrument for households size does not pass the specification tests in this model. The triple difference analysis does not show any statistically significant results in this case.

7.4 Sub sample analysis

I re-estimate my baseline school expenditure and school enrollment models in Eqns. 1, 2, 3 and 4 to find the treatment impact by changing the samples.

First, I estimate three sub-sample regressions for school enrollment analysis by varying age groups. Table 13, panels A and B present the difference-in-differences and triple differences results. I restrict the sample to children in the age group 5-9 years, 10-14 years and 15-17 years. I find consistent results with the baseline for the sub-sample of 5-9 years and 10-14 years. For both the age groups, the gender gap in enrollment reduces by 1.9 percentage points as a result of access to health insurance. I do not find any impact of RSBY for the sub sample 15-17 years. This could possibly be explained by lower marginal benefit of keeping older children in school than that of younger children.

Second, I conduct sub sample regressions of both school expenditure and enrollment models for rural and urban areas separately. Table 14 presents the results. From Panel I, I find RSBY to have a larger treatment effect on household's budget share of school expenditure in urban areas than rural from the DID model. The treatment effect is found to be 0.7 percentage points in the urban areas compared to 0.4 percentage points rise in the rural areas. Both effects are statistically significant. However, from the triple differences model, I do not find a statistically significant impact of RSBY in the urban areas. In contrast, RSBY has a treatment effect of 0.8 percentage points increase in the budget share of school expenditure for the poor households in the treated districts of rural areas. For the school enrollment model (see panel II), RSBY reduces the gender gap in enrollment by 1.5 to 2.4 percentage points in the rural areas, perhaps owing to the low levels of girls' enrollments to begin with. No such impact is found in the urban areas from either the DID or the DDD models. This speaks to the effectiveness of having access to such a health insurance in rural parts.

Table 15 presents the sub sample results for both expenditure and enrollment models estimated by caste categories. Panel I suggests that the treatment has a positive effect on school expenditure for the general category and other backward castes. This can be seen from the DID models. The DDD model also suggests a positive impact on the poor households in the treated districts belonging to other caste categories, apart from those in belonging to general castes. Panel II suggests that RSBY reduces the gender gap in enrollment to a large extent for the other backward castes (OBC) category. The magnitude of reduction in the gender gap for OBCs is economically large. This is confirmed by both the DID and DDD models. The triple differences model also suggests reduction in the gender gap of enrollments for the scheduled tribes and other castes. However, these results are not confirmed in the DID models.

8 Conclusion

Gender gap in schooling remains a concern for most policy makers and educationists. The UN Millennium Development Goals first enunciated in 2000, emphasized reducing gender gap in school that disadvantage girls (Grant and Behrman (2010), Nations (2015)). Such differences in education could potentially lead to further gender inequalities in income, work and social status. Given that women are significant contributors in the labour force in most developing countries, gender gaps can act as constraints on economic growth. In fact, investment in female education is widely regarded as essential by policymakers owing to the positive externalities associated with it, such as, better child health, household welfare, and lower population growth (Song et al., 2006, Alderman and King, 1998). Appropriate policy responses to reduce the gender gap thus require an understanding of its determinants. Little evidence exists on the impact of health insurance on school expenditure, in general, and on this gender gap, in particular, in India and my paper attempts to analyze this unexplored determinant.

Understanding the impact of shocks on education decisions of vulnerable households and the channels of this impact could help in designing safety nets and other policies to insulate investments in education from health shocks (Glick et al. (2016)). An undesired consequence of negative health shocks may be taking children out of school either to protect their health or to send them to work for additional income. These strategies can have undesired consequences in the long term for human capital accumulation of future generations and labor market opportunities.

From a policy perspective, it is not only interesting to see if a health insurance scheme has an unintended role to play on school expenditure decisions of households but also on parental response within household in terms of enrollments of boys versus girls. At the outset, it is not entirely obvious as to whether health insurance would benefit children's education or have a detrimental impact. Healthier children could either mean greater future economic returns from schooling or greater value as child labour. Such responses need to be considered when designing policies to remedy any disadvantages among children because parents can exacerbate or eliminate these effects by aiming at equitable child human capital formation within the family.

Although RSBY was implemented with the intention of reducing financial burden for the poor, I find that it has unintended positive consequences for children. First, I find that household expenditure on school increases as a result of access to RSBY. Second, I find that access to a health insurance systems provides additional resources to parents, in a society that depends largely on sons for support during old age, to not exclude their daughters from education opportunities. Robustness checks and sensitivity analyses support the validity of my results. This is evidence that health insurance protects the poor and also helps such households keep their children in school in the face of health shocks. This unintended benefit could help push households out of the vicious cycle of poor health in childhood leading to lesser education and hence lower incomes and health in adulthood. In addition, there is also a long-term positive effect of health insurance coverage on economic development, this effect being reinforced through the positive impact on school enrollments of girls.

References

Acharya, A., Vellakkal, S., Taylor, F., Masset, E., Satija, A., Burke, M., and Ebrahim, S. (2012). The impact of health insurance schemes for the informal sector in low-and middle-income countries: a systematic review. *The World Bank Research Observer*, 28(2):236–266.

Alam, S. A. (2015). Parental health shocks, child labor and educational outcomes: Evidence from tanzania. *Journal of health economics*, 44:161–175.

Alderman, H., Behrman, J. R., Ross, D. R., and Sabot, R. (1996). Decomposing the gender gap in cognitive skills in a poor rural economy. *Journal of Human Resources*, pages 229–254.

Alderman, H. and Gertler, P. (1997). Family resources and gender differences in human capital investments: The demand for children's medical care in pakistan. *Intrahousehold Resource Allocation in Developing Countries*, pages 231–48.

Alderman, H. and King, E. M. (1998). Gender differences in parental investment in education. *Structural Change and Economic Dynamics*, 9(4):453–468.

Almond, D., Li, H., and Meng, L. (2010). Son preference and early childhood investments in china. *manuscript, Tsinghua University.*

Azam, M. (2018). Does social health insurance reduce financial burden? panel data evidence from india. *World Development*, 102:1–17.

Barcellos, S. H., Carvalho, L. S., and Lleras-Muney, A. (2014). Child gender and parental investments in india: are boys and girls treated differently? *American Economic Journal: Applied Economics*, 6(1):157–189.

Bardhan, K. (1985). Women's work, welfare and status: Forces of tradition and change in india. *Economic and Political Weekly*, pages 2207–2220.

Baum, C. F. et al. (2013). Implementing new econometric tools in stata. In *Mexican Stata Users' Group Meetings*, volume 9.

Baum, C. F., Schaffer, M. E., Stillman, S., et al. (2007). Enhanced routines for instrumental variables/gmm estimation and testing. *Stata Journal*, 7(4):465–506.

Behrman, J. R. (1988). Intrahousehold allocation of nutrients in rural india: Are boys favored? do parents exhibit inequality aversion? *Oxford Economic Papers*, 40(1):32–54.

Behrman, J. R. and Knowles, J. C. (1999). Household income and child schooling in vietnam. *The World Bank Economic Review*, 13(2):211–256.

Bhalotra, S. and Attfield, C. (1998). Intrahousehold resource allocation in rural pakistan: a semiparametric analysis. *Journal of Applied Econometrics*, pages 463–480.

Bhalotra, S. and Rawlings, S. B. (2011). Intergenerational persistence in health in developing countries: The penalty of gender inequality? *Journal of Public Economics*, 95(3):286–299.

Bratti, M. and Mendola, M. (2014). Parental health and child schooling. *Journal of health economics*, 35:94–108.

Burbidge, J. B., Magee, L., and Robb, A. L. (1988). Alternative transformations to handle extreme values of the dependent variable. *Journal of the American Statistical Association*, 83(401):123–127.

Burgess, R. and Zhuang, J. (2000). Modernisation and son preference.

Chen, Y. and Jin, G. Z. (2012). Does health insurance coverage lead to better health and educational outcomes? evidence from rural china. *Journal of health economics*, 31(1):1–14.

Clark, S. (2000). Son preference and sex composition of children: Evidence from india. *Demography*, 37(1):95–108.

Clarke, D. (2017). Chidlren and their parents: A review of fertility and causality. *Journal of Economic Surveys*.

Colclough, C., Rose, P., and Tembon, M. (2000). Gender inequalities in primary schooling: The roles of poverty and adverse cultural practice. *International Journal of Educational Development*, 20(1):5–27.

Currie, J. and Gruber, J. (1996). Health insurance eligibility, utilization of medical care, and child health. *The Quarterly Journal of Economics*, 111(2):431–466.

Currie, J. and Moretti, E. (2007). Biology as destiny? short-and long-run determinants of intergenerational transmission of birth weight. *Journal of Labor economics*, 25(2):231–264.

Das, J. and Leino, J. (2011). Evaluating the rsby: lessons from an experimental information campaign. *Economic and Political Weekly*, pages 85–93.

Deaton, A. (1997). *The analysis of household surveys: a microeconometric approach to development policy*. World Bank Publications.

Devadasan, N., Seshadri, T., Trivedi, M., and Criel, B. (2013). Promoting universal financial protection: evidence from the rashtriya swasthya bima yojana (rsby) in gujarat, india. *Health Research Policy and Systems*, 11(1):29.

Dreze, J. and Kingdon, G. G. (2001). School participation in rural india. *Review of Development Economics*, 5(1):1–24.

Duraisamy, P. (1992). Gender, intrafamily allocation of resources and child schooling in south india.

Edmonds, E. V. (2006). Child labor and schooling responses to anticipated income in south africa. *Journal of development Economics*, 81(2):386–414.

Fabre, A. and Pallage, S. (2015). Child labor, idiosyncratic shocks, and social policy. *Journal of Macroeconomics*, 45:394–411.

Flores, G., Krishnakumar, J., O'donnell, O., and Van Doorslaer, E. (2008). Coping with health-care costs: implications for the measurement of catastrophic expenditures and poverty. *Health economics*, 17(12):1393–1412.

Garg, A. and Morduch, J. (1998). Sibling rivalry and the gender gap: Evidence from child health outcomes in ghana. *Journal of Population Economics*, 11(4):471–493.

Glick, P. J., Sahn, D. E., and Walker, T. F. (2016). Household shocks and education investments in madagascar. *Oxford Bulletin of Economics and Statistics*, 78(6):792–813.

Government of India, G. (2013). Rashtriya swasthya bima yojana.

Grant, M. J. and Behrman, J. R. (2010). Gender gaps in educational attainment in less developed countries. *Population and Development Review*, 36(1):71–89.

Haddad, W., Carnoy, M., and Rinaldi, R. (1984). 0. regel. 1990." education and development: Evidence for new priorities.". *World Bank Discussion Papers*, 95.

Hamoudi, A. A., Sachs, J. D., et al. (1999). Economic consequences of health status: a review of the evidence. Technical report, Center for International Development at Harvard University.

Hazarika, G. and Sarangi, S. (2008). Household access to microcredit and child work in rural malawi. *World Development*, 36(5):843–859.

Hill, M. A. and King, E. (1995). Women's education and economic well-being. *Feminist Economics*, 1(2):21–46.

Horrace, W. C. and Oaxaca, R. L. (2006). Results on the bias and inconsistency of ordinary least squares for the linear probability model. *Economics Letters*, 90(3):321–327.

Johnson, D. and Krishnaswamy, K. (2012). The impact of rsby on hospital utilization and out-of-pocket health expenditure. *Washington DC: World Bank Group*.

Kambhampati, U. S. and Pal, S. (2001). Role of parental literacy in explaining gender difference: Evidence from child schooling in india. *The European Journal of Development Research*, 13(2):97–119.

Karan, A., Selvaraj, S., and Mahal, A. (2014). Moving to universal coverage? trends in the burden of out-of-pocket payments for health care across social groups in india, 1999–2000 to 2011–12. *PloS one*, 9(8):e105162.

Karan, A., Yip, W., and Mahal, A. (2017). Extending health insurance to the poor in india: An impact evaluation of rashtriya swasthya bima yojana on out of pocket spending for healthcare. *Social Science & Medicine*, 181:83–92.

Katz, E. G. (1995). Gender and trade within the household: observations from rural guatemala. *World Development*, 23(2):327–342.

Kingdon, G. G. (2005). Where has all the bias gone? detecting gender bias in the intrahousehold allocation of educational expenditure. *Economic Development and Cultural Change*, 53(2):409–451.

Kochar, A. (1995). Explaining household vulnerability to idiosyncratic income shocks. *The American Economic Review*, 85(2):159–164.

Lake, J. and Millimet, D. L. (2016). An empirical analysis of trade-related redistribution and the political viability of free trade. *Journal of International Economics*, 99:156–178.

Liu, H. and Zhao, Z. (2014). Does health insurance matter? evidence from china's urban resident basic medical insurance. *Journal of Comparative Economics*, 42(4):1007–1020.

Liu, K. (2016). Insuring against health shocks: Health insurance and household choices. *Journal of health economics*, 46:16–32.

Mohanan, M. (2013). Causal effects of health shocks on consumption and debt: quasi-experimental evidence from bus accident injuries. *Review of Economics and Statistics*, 95(2):673–681.

Morduch, J. (1999). Between the state and the market: Can informal insurance patch the safety net? *The World Bank Research Observer*, 14(2):187–207.

Mundlak, Y. (1978). On the pooling of time series and cross section data. *Econometrica: journal of the Econometric Society*, pages 69–85.

Nandi, A., Ashok, A., and Laxminarayan, R. (2013). The socioeconomic and institutional determinants of participation in India's health insurance scheme for the poor. *PloS one*, 8(6):e66296.

Nations, U. (2015). The millenium development goals report 2015. Technical report, United Nations.

Pal, S. (2004). How much of the gender difference in child school enrolment can be explained? evidence from rural india. *Bulletin of Economic Research*, 56(2):133–158.

Palacios, R., Das, J., and Sun, C. (2011). India's health insurance scheme for the poor: Evidence from the early experience of the rashtriya swasthya bima yojana. *New Delhi: Center for Policy Research*.

Papke, L. E. and Wooldridge, J. M. (2008). Panel data methods for fractional response variables with an application to test pass rates. *Journal of Econometrics*, 145(1):121–133.

Rajasekhar, D., Berg, E., Ghatak, M., Manjula, R., and Roy, S. (2011). Implementing health insurance: the rollout of rashtriya swasthya bima yojana in karnataka. *Economic and Political Weekly*, pages 56–63.

Ravi, S. and Bergkvist, S. (2015). Are publicly financed health insurance schemes working in india? *India Policy Forum*, 11(1):158–192.

Raza, W., van de Poel, E., and Panda, P. (2016). Analyses of enrolment, dropout and effectiveness of rsby in northern rural india.

Rosenzweig, M. R. and Schultz, T. P. (1982). Market opportunities, genetic endowments, and intrafamily resource distribution: Child survival in rural india. *The American Economic Review*, 72(4):803–815.

Sauerborn, R., Adams, A., and Hien, M. (1996). Household strategies to cope with the economic costs of illness. *Social science & medicine*, 43(3):291–301.

Schultz, T. P. (2002). Why governments should invest more to educate girls. *World Development*, 30(2):207–225.

Sen, A. and Sengupta, S. (1983). Malnutrition of rural children and the sex bias. *Economic and political weekly*, pages 855–864.

Skoufias, E. (1993). Labor market opportunities and intrafamily time allocation in rural households in south asia. *Journal of Development Economics*, 40(2):277–310.

Song, L., Appleton, S., and Knight, J. (2006). Why do girls in rural china have lower school enrollment? *World Development*, 34(9):1639–1653.

Sun, A. and Yao, Y. (2010). Health shocks and children's school attainments in rural china. *Economics of Education Review*, 29(3):375–382.

Swarup, A. and Jain, N. (2011). Rashtriya swasthya bima yojana. *Innovative*, page 259.

United Nations, U. (2017). Gender-biased sex selection.

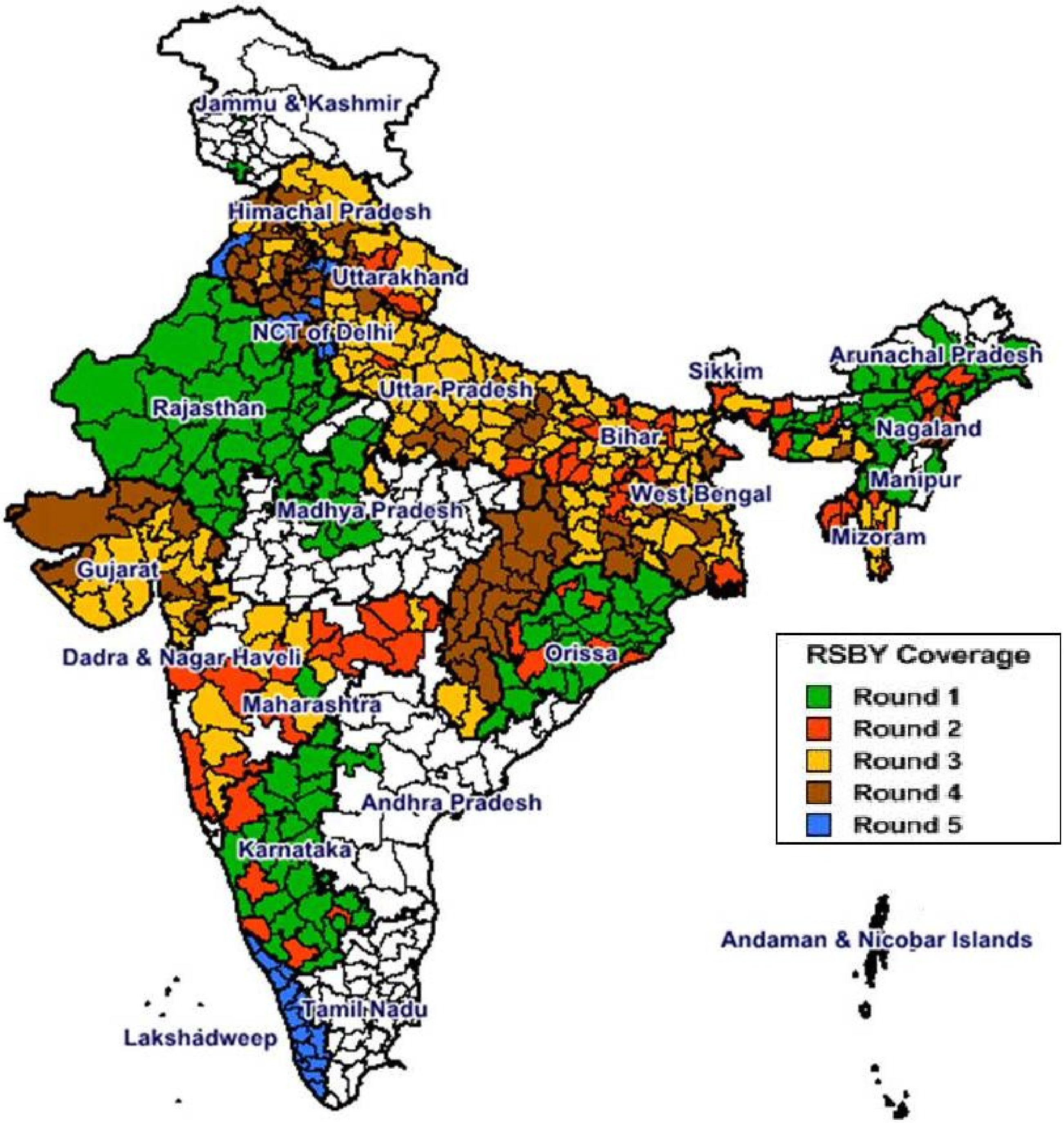
Virk, A. and Atun, R. (2015). Towards universal health coverage in india: a historical examination of the genesis of rashtriya swasthya bima yojana–the health insurance scheme for low-income groups. *Public health*, 129(6):810–817.

Wagstaff, A., Lindelow, M., Jun, G., Ling, X., and Juncheng, Q. (2009). Extending health insurance to the rural population: an impact evaluation of china's new cooperative medical scheme. *Journal of health economics*, 28(1):1–19.

WHO (2015). Tracking universal health coverage, first global monitoring report. Technical report, World Health Organization and The World Bank.

Wooldridge, J. M. (2015). Control function methods in applied econometrics. *Journal of Human Resources*, 50(2):420–445.

Wooldridge, J. M. et al. (2011). Fractional response models with endogeneous explanatory variables and heterogeneity. In *CHI11 Stata Conference*, number 12. Stata Users Group.



Source: www.rsby.gov.in

Table 1. Summary Statistics - Household Level

	Control Districts				Treated Districts			
	2004-05		2011-12		2004-05		2011-12	
Variables	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**
School expenditure	59.55	129.67	137.08	272.57	104.13	218.84	162.60	288.19
Total consumption expenditure	3527.21	3315.30	7719.33	7977.05	4022.48	3917.97	7631.89	7318.71
Age	45.51	13.13	46.27	13.20	46.82	13.33	47.08	13.60
Household size	6.38	2.95	5.56	2.14	6.71	3.09	5.85	2.31
Urban (1 = yes)	0.30	0.46	0.31	0.46	0.27	0.44	0.29	0.45
Other Backward Castes (1 = yes)	0.42	0.49	0.21	0.41	0.40	0.49	0.21	0.41
Scheduled Caste (1 = yes)	0.21	0.41	0.43	0.49	0.22	0.42	0.41	0.49
Scheduled Tribe (1 = yes)	0.10	0.30	0.19	0.40	0.08	0.27	0.24	0.42
Other castes (1 = yes)	0.24	0.43	0.14	0.34	0.24	0.43	0.09	0.29
Muslim (1 = yes)	0.09	0.28	0.09	0.28	0.14	0.35	0.16	0.36
Christian (1 = yes)	0.02	0.12	0.03	0.17	0.03	0.16	0.02	0.14
Sikh or Buddhist (1 = yes)	0.03	0.16	0.02	0.14	0.04	0.19	0.03	0.18
Other religion (1 = yes)	0.01	0.09	0.00	0.06	0.01	0.11	0.01	0.08
HH Head - literate (1 = yes)	0.64	0.48	0.70	0.46	0.63	0.48	0.66	0.47
HH Head - knows english (1 = yes)	0.15	0.35	0.16	0.36	0.18	0.38	0.16	0.37
HH Head - ever attended school (1 = yes)	0.65	0.48	0.65	0.48	0.63	0.48	0.61	0.49
Male with primary education (1 = yes)	0.15	0.35	0.15	0.35	0.14	0.35	0.15	0.36
Male with secondary education (1 = yes)	0.28	0.45	0.38	0.49	0.26	0.44	0.37	0.48
Male with senior sec. education (1 = yes)	0.06	0.24	0.14	0.35	0.06	0.24	0.12	0.33
Male with college education (1 = yes)	0.04	0.20	0.13	0.33	0.05	0.22	0.11	0.31
Female with primary education (1 = yes)	0.17	0.37	0.16	0.36	0.16	0.37	0.15	0.36
Female with secondary education (1 = yes)	0.37	0.48	0.35	0.48	0.38	0.48	0.30	0.46
Female with senior sec. education (1 = yes)	0.12	0.33	0.10	0.30	0.10	0.30	0.09	0.29
Female with college education (1 = yes)	0.09	0.29	0.07	0.26	0.10	0.30	0.08	0.26
Gender of the head (1 = male)	0.94	0.23	0.91	0.29	0.92	0.28	0.87	0.34
# of married males	1.45	0.86	1.30	0.70	1.45	0.88	1.26	0.73
# of married females	1.48	0.86	1.34	0.71	1.53	0.90	1.38	0.73
Proportion of children	0.39	0.15	0.37	0.14	0.39	0.15	0.38	0.16
HH has a bank account (1=yes)	0.34	0.48	0.34	0.47	0.36	0.48	0.34	0.47
HH has a Kisan credit card (1=yes)	0.04	0.20	0.06	0.24	0.05	0.21	0.05	0.23
HH has a credit card (1=yes)	0.01	0.11	0.02	0.15	0.01	0.12	0.03	0.16

Notes: Sample is resticted to households where age of the head of the household is between 18 to 90 years. The table shows the summary statistics in the control districts and treatment districts in 2004-05 and 2011-12 for household level. Dummy variables containing information about education levels, demography, bank information, caste and religion of the household are included. Muslim takes value 1 if the household is Muslim, 0 otherwise. Christian = 1 if the household is Christian, 0 otherwise. Sikh = 1 if the household is Sikh, 0 otherwise. Other religion = 1 if the the household falls under any of the other categories like Jainism, Buddhism, Zoroastrianism, and others, 0 otherwise. ST = 1 if the household is scheduled tribe, 0 otherwise. SC = 1 if the household is scheduled caste, 0 otherwise. OBC =1 if the household belongs to other backward castes, 0 otherwise. Other religion =1 if the household belongs to other other castes or general category, 0 otherwise.

Table 2. Summary Statistics - Individual Level

Variables	Control Districts 2004-05 Mean	SD	Control Districts 2011-12 Mean	SD	Treated Districts 2004-04 Mean	SD	Treated Districts 2011-12 Mean	SD
Enrolled (1 = yes)	0.795	0.404	0.886	0.317	0.758	0.429	0.865	0.342
Gender (1 = boy)	0.522	0.500	0.535	0.499	0.525	0.499	0.525	0.499
Household size	6.919	3.213	7.153	3.219	7.100	3.149	7.377	3.347
Scholarship from school (1 = yes)	0.074	0.261	0.365	0.481	0.089	0.285	0.350	0.477
Age	10.968	3.683	11.469	3.609	10.876	3.712	11.321	3.634
Benefits from school (1 = yes)	0.487	0.500	0.547	0.498	0.344	0.475	0.434	0.496
Urban (1 = yes)	0.274	0.446	0.251	0.434	0.268	0.443	0.287	0.452
OBC (1 = yes)	0.355	0.479	0.257	0.437	0.384	0.486	0.228	0.420
SC (1 = yes)	0.170	0.376	0.374	0.484	0.234	0.423	0.411	0.492
ST (1 = yes)	0.159	0.366	0.162	0.369	0.054	0.226	0.244	0.429
Other castes (1 = yes)	0.272	0.445	0.162	0.369	0.268	0.443	0.062	0.241
Muslim (1 = yes)	0.116	0.320	0.102	0.303	0.176	0.380	0.181	0.385
Christian (1 = yes)	0.038	0.190	0.028	0.164	0.017	0.131	0.016	0.125
Sikh & Buddhist (1 = yes)	0.033	0.177	0.026	0.159	0.043	0.203	0.038	0.191
Other religion (1 = yes)	0.019	0.135	0.006	0.075	0.011	0.103	0.004	0.066
Father with primary education (1 = yes)	0.167	0.373	0.159	0.366	0.148	0.355	0.150	0.357
Father with secondary education (1 = yes)	0.251	0.434	0.372	0.483	0.224	0.417	0.353	0.478
Father with senior sec. education (1 = yes)	0.054	0.225	0.129	0.335	0.042	0.201	0.107	0.309
Father with college educatiom (1 = yes)	0.032	0.175	0.107	0.309	0.036	0.185	0.089	0.285
Mother with primary education (1 = yes)	0.178	0.382	0.170	0.375	0.161	0.367	0.154	0.361
Mother with secondary education (1 = yes)	0.371	0.483	0.309	0.462	0.370	0.483	0.257	0.437
Mother with senior sec. education (1 = yes)	0.111	0.314	0.078	0.268	0.085	0.279	0.067	0.249
Mother with college education (1 = yes)	0.084	0.277	0.051	0.220	0.082	0.275	0.054	0.226

Notes: Sample is resticted to households where children between the age group 5 to 18 years. The table shows summary statistics in the control districts and treatment districts in 2004-05 and 2011-12 for the individual level data. Dummy variables containing information about gender, school facilities, parental education levels, age, caste and religion of the individuals are included. Muslim takes value 1 if individual is Muslim, 0 otherwise. Christian = 1 if individual is Christian, 0 otherwise. Sikh = 1 if individual is Sikh, 0 otherwise. Other religion = 1 if the individual falls under any of the other categories like Jainism, Buddhism, Zoroastrianism, and others, 0 otherwise. ST = 1 if individual's caste is scheduled tribe, 0 otherwise. SC = 1 if individual's caste is scheduled caste, 0 otherwise. OBC=1 if individual belongs to other bcakward caste, 0 otherwise. Other religion =1 if individual belongs to other castes or general category, 0 otherwise.

Table 3. Impact of RSBY on household school expenditure

	Panel A. DID			Panel B. DDD		
	(1) School expd. Budget Share	(2) Log School expd. Levels	(3) Log Total consumption expd. Levels	(4) School expd. Budget Share	(5) Log School expd. Levels	(6) Log Total consumption expd. Levels
RSBY*Post	0.005***	0.302***	0.077***	-0.003*	-0.120	-0.015
	(0.001)	(0.127)	(0.014)	(0.002)	(0.257)	(0.074)
Low Income (=1 for bottom 70%)				-0.047**	-0.829*	-0.224***
				(0.024)	(0.499)	(0.084)
RSBY*Post*Low Income				0.007***	0.422***	0.098
				(0.001)	(0.188)	(0.086)
Underidentification test	p=0.000	p=0.000	p=0.000	p=0.001	p=0.001	p=0.001
Weak-identification test						
Kleigbergen Paap rk Wald F statistic	11.646	11.607	11.574	6.580	6.543	6.553
Endogeneity test	p=0.010	p=0.000	p=0.025	p=0.010	p=0.002	p=0.027
Other Controls	Y	Y	Y	Y	Y	Y
District fixed effects	Y	Y	Y	Y	Y	Y
Time fixed effects	Y	Y	Y	Y	Y	Y
District*Income fixed effects				Y	Y	Y
Time*Income fixed effects				Y	Y	Y
N	47421	47421	47421	47421	47421	47421

* p<0.10, ** p<0.05, *** p<0.01. The sample is restricted to HH with children and where age of the head is between 18 to 90 years. Panel A and B provide the DID and DDD results respectively. Estimation is using IV approach Col(1) and (4): dependent variable is budget share of household's school expenditure (school expenditure/total consumption expenditure). Col. (2) and (5) : dependent variable is the inverse hyperbolic sine transformation of school expenditure in levels. Col. (3) and (6): dependent variable is the inverse hyperbolic sine transformation of total consumption expenditure in levels. Additional controls include: RSBY = 1 if the district was exposed to RSBY & 0 otherwise, dummy for Low Income =1 if HH does not belong to top 30% and 0 otherwise (for DDD), HH size (instrumented by gender of the first child), highest education degrees of male and female members, indicators for religion of HH, indicators for caste of HH, dummy for urban areas, number of married men in the HH, number of married women in the HH, proportion of children, teens and adults, indicator for if HOH is married, dummy for if the HH has a bank account, dummy for if the HH has a farmer credit card, district fixed effects, time fixed effects, district by income fixed effects (for DDD), time by income fixed effect (for DDD). Standard errors reported are clustered standard errors.

Table 4. Impact of RSBY on child school enrollment

	Panel A. DID		Panel B. DDD	
	(1) Enrollment	(2) Enrollment with gender differential	(3) Enrollment	(4) Enrollment with gender differential
RSBY*Post	0.017***	0.027***	-0.024	-0.023
	(0.005)	(0.006)	(0.017)	(0.017)
Boy	0.053***	0.060***	0.054***	0.055***
	(0.004)	(0.004)	(0.004)	(0.004)
Low Income (=1 for bottom 70%)			-3.328	-3.576
			(9.608)	(9.577)
RSBY*Post*Boy		-0.019***		
		(0.005)		
RSBY*Post*Low Income			0.042**	0.046**
			(0.019)	(0.020)
RSBY*Post*Low Income*Boy				-0.009***
				(0.001)
Underidentification test	p=0.000	p=0.000	p=0.000	p=0.000
Weak-identification test				
Kleigbergen Paap rk Wald F statistic	43.434	44.022	42.242	42.46
Endogeneity test	p=0.509	p = 0.570	p=0.401	p=0.414
Other Controls	Y	Y	Y	Y
District Fixed Effects	Y	Y	Y	Y
Time Fixed Effects	Y	Y	Y	Y
District*Income Fixed Effects			Y	Y
Time*Income Fixed Effects			Y	Y
N	83221	83221	83221	83221

* $p<0.10$, ** $p<0.05$, *** $p<0.01$. The sample is restricted to children above the age of 5 and below the age of 18. Panel A and B provide the DID and DDD results respectively. Estimation is using a LPM. Dependent variable is school enrollment of a child in a household. Additional controls include RSBY = 1 if the district was exposed to RSBY & 0 otherwise, dummy for Low Income =1 if HH does not belong to top 30% and 0 otherwise (for DDD), a gender dummy = 1 for a boy and 0 for a girl, RSBY, HH size, parental education characteristics, indicators for religion of HH, indicators for caste of HH, dummy for urban areas, school facilities and scholarships offered, district FE, time FE, district by income fixed effects (for DDD), time by income fixed effects (for DDD). HH size is instrumented by the gender of the first child. Standard errors reported are clustered standard errors.

Table 5. Robustness: Impact of RSBY on household school expenditure - Instrumental variable approach

	Panel A. DID		Panel B. DDD	
	(1)	(2)	(3)	(4)
	Budget Share	Log School expd. Levels	Budget Share	Log School expd. Levels
RSBY*Post	0.005***	0.080***	-0.003*	-0.105
	(0.001)	(0.014)	(0.002)	(0.169)
Low Income (=1 for bottom 70%)			-0.047**	-0.384**
			(0.024)	(0.169)
RSBY*Post*Low Income			0.007***	0.187***
			(0.001)	(0.088)
Underidentification test	p=0.000	p=0.000	p=0.001	p=0.013
Weak-identification test				
Kleigbergen Paap rk Wald F statistic	11.646	17.143	6.580	12.379
Endogeneity test	p=0.010	p=0.031	p=0.010	p=0.038
Other Controls	Y	Y	Y	Y
Total Consumption Expenditure	N	Y	N	Y
District fixed effects	Y	Y	Y	Y
Time fixed effects	Y	Y	Y	Y
District*Income fixed effects			Y	Y
Time*Income fixed effects			Y	Y
N	47421	47421	47421	47421

* p<0.10, ** p<0.05, *** p<0.01. The sample is restricted to HH with children and where age of the head is between 18 to 90 years. Panel A and B provide the DID and DDD results respectively. Col. (1) & (3) repeat the baseline IV result. Dependent variable is HH budget share of school expenditure. (2) & (4) are estimated via IV approach. Dependent variable is the inverse hyperbolic sine transformation of HH expenditure on school in levels. Total consumption expenditure is added as a regressor and instrument used for it is HH assets. Additional control in all regressions are RSBY = 1 if the district was exposed to RSBY & 0 otherwise, dummy for Low Income =1 if HH does not belong to top 30% and 0 otherwise (for DDD), HH size (instrumented by gender of the first child), highest education degrees of male and female members, indicators for religion of HH, indicators for caste of HH, dummy for urban areas, number of married men in the HH, number of married women in the HH, proportion of children, teens and adults, indicator for if HOH is married, dummy for if the HH has a bank account, dummy for if the HH has a farmer credit card, district fixed effects, time fixed effects, district by income fixed effects (for DDD), time by income fixed effects (for DDD). Standard errors reported are clustered standard errors.

Table 6. Robustness: Impact of RSBY on household school expenditure - Fractional logit estimation

	Panel A. DID		Panel B. DDD	
	(1)	(2)	(3)	(4)
	IV with FE	CRE FracLogit (control function)	IV with FE	CRE FracLogit (control function)
RSBY*Post	0.005***	0.029	-0.003*	0.008
	(0.001)	(0.064)	(0.002)	(0.068)
Low Income (=1 for bottom 70%)			-0.047**	0.069
			(0.024)	(0.091)
RSBY*Post*Low Income			0.007***	0.067
			(0.001)	(0.043)
Marginal Effect of RSBY:		0.001		
		(0.001)		
Households that belong to bottom 70%				0.002**
				(0.001)
Households that belong to top 30%				0.000
				(0.002)
Underidentification test	p=0.000		p=0.001	
Weak-identification test				
Kleigbergen Paap rk Wald F statistic	11.646		6.580	
Endogeneity test	p=0.010		p=0.010	
Other Controls	Y	Y	Y	Y
District fixed effects	Y	N	Y	N
Correlated random effects	N	Y	N	Y
Time fixed effects	Y	Y	Y	Y
Time*Income fixed effects			Y	Y
N	47421	47421	47421	47421

* p<0.10, ** p<0.05, *** p<0.01. The sample is restricted to HH with children and where age of the head is between 18 to 90 years. Panel A and B provide the DID and DDD results respectively. Col. (1) and (3) repeat the baseline IV with fixed effects results. Col. (2) and (4) via a fractional logit model with correlated random effects using a control function approach. Dependent variable in all specifications is household's budget share of school expenditure. Additional controls in all specifications include : RSBY = 1 if the district was exposed to RSBY & 0 otherwise, dummy for Low Income =1 if HH does not belong to top 30% and 0 otherwise (for DDD), HH size (instrumented by gender of the first child), highest education degrees of male and female members, indicators for religion of HH, indicators for caste of HH, dummy for urban areas, number of married men in the HH, number of married women in the HH, proportion of children, teens and adults, indicator for if HOH is married, dummy for if the HH has a bank account, dummy for if the HH has a farmer credit card, district fixed effects, time fixed effects, district by income fixed effects (for DDD), time by income fixed effects (for DDD). Standard errors reported are clustered standard errors.

Table 7. Robustness: Impact of RSBY on household school expenditure - Panel analysis

	Panel A. DID		Panel B. DDD	
	(1) Repeated Cross-Section	(2) Panel	(3) Repeated Cross-Section	(4) Panel
RSBY*Post	0.005***	0.003***	-0.003*	-0.001
	(0.001)	(0.001)	(0.002)	(0.002)
Low Income (=1 for bottom 70%)			-0.047**	-0.776*
			(0.024)	(0.445)
RSBY*Post*Low Income			0.007***	0.004**
			(0.001)	(0.002)
Other Controls	Y	Y	Y	Y
District Fixed Effects	Y	N	Y	N
Household Fixed Effects	N	Y	N	Y
Time Fixed Effects	Y	Y	Y	Y
District*Income Fixed Effects			Y	Y
Time*Income Fixed Effects			Y	Y
N	47421	45676	47421	45676

* p<0.10, ** p<0.05, *** p<0.01. The sample is restricted to HH with children and where age of the head is between 18 to 90 years. Panel A and B provide the DID and DDD results respectively. Dependent variable in all specifications is budget share of household's school expenditure. Col. (1) and (3) repeat the baseline IV with FE results. Data is treated in baseline as a repeated cross-section. Col (2) & (4) are estimated treating data as a panel data using IV with HH FE. Additional controls include: RSBY = 1 if the HH in the district was exposed to RSBY & 0 otherwise, dummy for Low Income =1 if HH does not belong to top 30% and 0 otherwise (for DDD), HH size (instrumented by the gender of the first child), highest education degrees of male and female members, indicators for religion of HH, indicators for caste of HH, dummy for urban areas, number of married men in the HH, number of married women in the HH, proportion of children, teens and adults, indicator for if HOH is married, dummy for if the HH has a bank account, dummy for if the HH has a farmer credit card, district fixed effects, HH fixed effects and time fixed effects. Standard errors reported are clustered standard errors.

Table 8. Robustness: Impact of RSBY on child school enrollment - Probit with correlated random effects

	Panel A. DID			Panel B. DDD		
	(1)	(2)	(3)	(4)	(5)	(6)
	LPM with FE	LPM with CRE	CRE Probit	LPM with FE	LPM with CRE	CRE Probit
RSBY*Post	0.027***	0.017***	0.126**	-0.023	-0.023	0.191***
	(0.006)	(0.005)	(0.060)	(0.017)	(0.017)	(0.073)
Boy	0.060***	0.059***	0.293***	0.055***	0.055***	0.253***
	(0.004)	(0.004)	(0.021)	(0.004)	(0.004)	(0.045)
RSBY*Post*Boy	-0.019***	-0.018***	-0.032*			
	(0.005)	(0.005)	(0.014)			
Low Income (=1 for bottom 70%)				-0.042	-0.043	-0.123
				(0.023)	(0.023)	(0.187)
RSBY*Post*Low Income				0.046**	0.046**	-0.151
				(0.020)	(0.020)	(0.103)
RSBY*Post*Low Income*Boy				-0.009***	-0.009***	0.038
				(0.001)	(0.001)	(0.075)
Marginal Effects of RSBY:						
Boy			0.094*			0.026
			(0.055)			(0.068)
Girl			0.126**			0.040
			(0.060)			(0.073)
Underidentification test	p=0.000	p=0.000		p=0.000	p=0.000	
Weak-identification test						
Kleigbergen Paap rk Wald F statistic	44.022	48.396		42.46	50.666	
Endogeneity test	p = 0.570	p=0.546		p=0.414	p=0.423	
Other Controls	Y	Y	Y	Y	Y	Y
District Fixed Effects	Y	N	N	Y	N	N
Correlated Random Effects	N	Y	Y	Y	Y	Y
Time Fixed Effects	Y	Y	Y	N	Y	Y
District*Income Fixed Effects				Y	Y	Y
Time*Income Fixed Effects				Y	Y	Y
N	83221	83221	83221	83221	83221	83221

* p<0.10, ** p<0.05, *** p<0.01. The sample is restricted to children above the age of 5 and below the age of 18. Panel. A provides the DID results and Panel. B provides the DDD results. Col. (1) and (4) are estimated via LPM with FE. Col. (2) and (5) are estimated via LPM with correlated random effects. Col. (3) and (6) are estimated via IV probit model with correlated random effects. Dependent variable in all specifications is school enrollment of a child in a household in a district at a particular point in time. Additional controls include: gender dummy = 1 for a boy and 0 for a girl, RSBY = 1 if the district was exposed to RSBY and 0 otherwise, Low Income dummy =1 if HH does not belong to top 30% and 0 otherwise (for DDD), HH size, parental education characteristics, indicators for religion of HH, indicators for caste of HH, dummy for urban areas, school facilities and scholarships offered, district fixed effects, time fixed effects, district by income fixed effects (for DDD) and time by income fixed effects (for DDD). HH size is instrumented by the gender of the first child. Means of all controls at district level have been included in (2), (3), (5) and (6). Standard errors reported are clustered standard errors.

Table 9. Robustness: Impact of RSBY on child school enrollment - Instrumental variable approach

	Panel A. DID			Panel B. DDD		
	(1) LPM with FE, IV	(2) LPM with FE	(3) LPM with FE	(4) LPM with FE, IV	(5) LPM with FE	(6) LPM with FE
RSBY*Post	0.027***	0.028**	0.028**	-0.023	-0.012	-0.016
	(0.006)	(0.011)	(0.011)	(0.017)	(0.018)	(0.019)
Boy	0.060***	0.058***	0.059***	0.055***	0.053***	0.054***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.003)	(0.004)
RSBY*Post*Boy	-0.019***	-0.018***	-0.018***			
	(0.005)	(0.006)	(0.006)			
Low Income (=1 for bottom 70%)				-0.042	-0.059***	-0.052***
				(0.023)	(0.006)	(0.006)
RSBY*Post*Low Income				0.046**	0.034	0.038*
				(0.020)	(0.022)	(0.022)
RSBY*Post*Low Income*Boy				-0.009***	-0.008***	-0.008***
				(0.001)	(0.001)	(0.001)
Other Controls	Y	Y	Y	Y	Y	Y
Household Size	Y	Y	N	Y	Y	N
District Fixed Effects	Y	Y	Y	Y	Y	Y
Time Fixed Effects	Y	Y	Y	Y	Y	Y
District*Income Fixed Effects				Y	Y	Y
Time*Income Fixed Effects				Y	Y	Y
N	83221	83221	83221	83221	83221	83221

* p<0.10, ** p<0.05, *** p<0.01. The sample is restricted to children above the age of 5 and below the age of 18. Panel. A provides the DID results and Panel. B provides the DDD results. (1) and (4) are estimated via LPM with FE. (2) and (5) are estiamted via LPM with FE including HH size as a regressor but not instrumenting for it. (3) and (6) are estimated via LPM with FE excluding HH size as a regressor. Dependent variable is school enrollment of a child in a household in a district at a particular point in time. Additional controls included in each specification - gender dummy = 1 for a boy and 0 for a girl, RSBY = 1 if the district was exposed to RSBY and 0 otherwise, Low Income dummy =1 if HH does not belong to top 30% and 0 otherwise (for DDD), parental education characteristics, indicators for religion of HH, indicators for caste of HH, dummy for urban areas, school facilities and scholarships offered, district fixed effects, time fixed effects, district by income fixed effects, time by income fixed effects. HH size is included as a regressor and instrumented by gender of the first child in the family. Standard errors reported are clustered standard errors.

Table 10. Sensitivity analysis: Impact of RSBY on household school expenditure and child school enrollment - Variation in income categories

	Panel A. DID	Panel B. DDD	
	(1)	(2)	(3)
	Baseline	Top and Bottom 30%	Mid 40 and Top 30%
Panel I. School Expenditure			
RSBY*Post	0.005***	-0.001	-0.001
	(0.001)	(0.002)	(0.002)
Low Income		0.001	0.010
		(0.002)	(0.019)
RSBY*Post*Low Income		0.005*	0.002
		(0.003)	(0.003)
Other Controls	Y	Y	Y
District fixed effects	Y	Y	Y
Time fixed effects	Y	Y	Y
District*Income fixed effects		Y	Y
Time*Income fixed effects		Y	Y
N	47421	27592	32835
Panel II. School Enrollment			
RSBY*Post	0.027***	-0.024**	-0.040***
	(0.006)	(0.012)	(0.008)
Boy	0.060***	0.057***	0.040***
	(0.004)	(0.006)	(0.004)
Low Income		0.056	-0.086
		(0.158)	(0.092)
RSBY*Post*Boy	-0.019***		
	(0.005)		
RSBY*Post*Low Income		0.063*	0.049***
		(0.033)	(0.011)
RSBY*Post*Low Income*Boy		-0.002*	0.001
		(0.001)	(0.006)
Underidentification test	p=0.000	p=0.001	p=0.000
Weak-identification test			
Kleigbergen Paap rk Wald F statistic	44.022	16.721	39.832
Endogeneity test	p = 0.570	p=0.529	p = 0.947
Other Controls	Y	Y	Y
District Fixed Effects	Y	Y	Y
Time Fixed Effects	Y	Y	Y
District*Income Fixed Effects		Y	Y
Time*Income Fixed Effects		Y	Y
N	83221	47876	57884

* p<0.10, ** p<0.05, *** p<0.01. The sample is restricted to HH with children and where age of the head is between 18 to 90 years. Panel A and B provide the DID and DDD results respectively. Col (1) repeats the baseline IV results. Col (2) provides DDD results where sample is restricted to top and bottom 30% of income distribution. Middle 40% is dropped. Col. (3) provides the DDD results where sample is restricted to middle 40% and top 30% of income distribution. Bottom 30% is dropped. Dependent variable in all specifications is budget share of household's school expenditure. Additional controls include : RSBY = 1 if the district was exposed to RSBY & 0 otherwise, Low Income dummy =1 if HH belongs to bottom 30% and 0 if HH belongs to top 30% (for Col (2)), Low Income dummy =1 if HH belongs to middle 40% and 0 if belongs to top 30% (for Col. (3)), HH size (instrumented by gender of the first child), highest education degrees of male and female members, indicators for religion of HH, indicators for caste of HH, dummy for urban areas, number of married men in the HH, number of married women in the HH, proportion of children, teens and adults, indicator for if HOH is married, dummy for if the HH has a bank account, dummy for if the HH has a farmer credit card, district fixed effects, time fixed effects, districy by income fixed effects (for DDD), time by income fixed effects (for DDD). Standard errors reported are clustered standard errors.

Table 11. Sensitivity analysis: Impact of RSBY on household school expenditure and child school enrollment - Variation by intensity of treatment

	Panel A. DID	Panel B. DDD
	(1)	(2)
Panel I. School Expenditure		
RSBY*Post	0.000	0.001
	(0.001)	(0.001)
RSBY*Post*Intensity	0.005***	
	(0.001)	
RSBY*Post*Low Income		0.001
		(0.002)
RSBY*Post*Low Income*Intensity		0.004
		(0.003)
Underidentification test	p=0.004	p=0.000
Weak-identification test		
Kleigbergen Paap rk Wald F statistic	12.291	12.147
Endogeneity test	p=0.012	p=0.009
Other Controls	Y	Y
District fixed effects	Y	Y
Time fixed effects	Y	Y
District*Income fixed effects		Y
Time*Income fixed effects		Y
N	37885	37885
Panel II. School Enrollment		
Boy	0.060***	0.056***
	(0.004)	(0.004)
RSBY*Post	0.028***	-0.023
	(0.006)	(0.018)
RSBY*Post*Boy	-0.011**	
	(0.005)	
RSBY*Post*Intensity	0.000	
	(0.004)	
RSBY*Post*Intensity*Boy	-0.022***	
	(0.005)	
RSBY*Post*Low Income		0.048**
		(0.021)
RSBY*Post*Low Income*Boy		0.001
		(0.006)
RSBY*Post*Low Income*Intensity		-0.003
		(0.006)
RSBY*Post*Low Income*Intensity*Boy		-0.027***
		(0.007)
Underidentification test	p=0.000	p=0.000
Weak-identification test		
Kleigbergen Paap rk Wald F statistic	43.943	44.56
Endogeneity test	p=0.588	p=0.474
Other Controls	Y	Y
District Fixed Effects	Y	Y
Time Fixed Effects	Y	Y
District*Income Fixed Effects		Y
Time*Income Fixed Effects		Y
N	83221	83221

p<0.10, ** p<0.05, *** p<0.01. The sample is restricted to HH with children and where age of the head is between 18 to 90 years. Panel A and B provide the DID and DDD results respectively. Panel I provides the school expenditure results and panel II, the school enrollment. Dependent variable in Panel I budget share of household's school expenditure. Dependent variable in Panel II is school enrollment of a child in a household in a district at a particular point in time. Additional controls include in panel A: RSBY = 1 if the district was exposed to RSBY & 0 otherwise, Low Income dummy =1 if HH belongs to bottom 30% and 0 otherwise, discrete indicator variable for intensity depending on duration of treatment, relevant two way and three way interaction with intensity, HH size (instrumented by gender of the first child), highest education degrees of male and female members, indicators for religion of HH, indicators for caste of HH, dummy for urban areas, number of married men in the HH, number of married women in the HH, proportion of children, teens and adults, indicator for if HOH is married, dummy for if the HH has a bank account, dummy for if the HH has a farmer credit card, district fixed effects, time fixed effects, district by income fixed effects (for DDD), time by income fixed effects (for DDD). Other additional controls in panel B: parental education characteristics, school facilities and scholarships offered. Standard errors reported are clustered standard errors.

Table 12. Sensitivity analysis: Impact of RSBY on household school expenditure - Variation in take-up by district

	Panel A		Panel B
	(1)	(2)	(3)
	DID	DID with district enrollment	DDD with district enrollment
RSBY*Post	0.007	0.003	-0.098
	(0.014)	(0.015)	(0.177)
RSBY*Post*DistrictEnrollment		0.006*	
		(0.004)	
RSBY*Post*Low Income			0.106
			(0.202)
RSBY*Post*Low Income*DistrictEnrollment			0.001
			(0.010)
Underidentification test	p=0.285	p=0.143	p=0.124
Weak-identification test			
Kleigbergen Paap rk Wald F statistic	1.111	2.116	2.203
Endogeneity test	p=0.322	p=0.355	p=0.313
Other Controls		Y	Y
District fixed effects		Y	Y
Time fixed effects		Y	Y
District*Income fixed effects			Y
Time*Income fixed effects			Y
N	15265	15265	15265

p<0.10, ** p<0.05, *** p<0.01. The sample is restricted to HH with children and where age of the head is between 18 to 90 years. Panel A and B provide the DID and DDD results respectively. Dependent variable is the budget share of household's school expenditure. Additional controls include : RSBY = 1 if the district was exposed to RSBY & 0 otherwise, District enrollment rate(= enrolled targeted households/total eligible households), Low Income dummy =1 if HH belongs to bottom 30% and 0 otherwise, HH size (instrumented by gender of the first child), highest education degrees of male and female members, indicators for religion of HH, indicators for caste of HH, dummy for urban areas, number of married men in the HH, number of married women in the HH, proportion of children, teens and adults, indicator for if HOH is married, dummy for if the HH has a bank account, dummy for if the HH has a farmer credit card, district fixed effects, time fixed effects, district by income fixed effects (for DDD), time by income fixed effects (for DDD). Standard errors reported are clustered standard errors.

Table 13. Sensitivity analysis: Impact of RSBY on child school enrollment - Variation in age groups

	Panel A. DID			Panel B. DDD		
	5-9 years	10-14 years	15-17 years	5-9 years	10-14 years	15-17 years
RSBY*Post	0.031***	0.035***	0.017	-0.022	-0.041***	0.070***
	(0.009)	(0.009)	(0.014)	(0.021)	(0.011)	(0.016)
Boy	0.043***	0.061***	0.081***	0.039***	0.058***	-0.097***
	(0.005)	(0.008)	(0.017)	(0.004)	(0.008)	(0.022)
RSBY*Post*Boy	-0.019***	-0.019***	-0.012			
	(0.007)	(0.007)	(0.012)			
Low Income (=1 for bottom 70%)				0.044	-0.085	-0.154
				(0.140)	(0.124)	(0.327)
RSBY*Post*Low Income				0.037	0.069***	0.111***
				(0.029)	(0.020)	(0.032)
RSBY*Post*Low Income*Boy				-0.009***	-0.016*	0.019
				(0.001)	(0.008)	(0.014)
Underidentification test	p=0.000	p=0.000	p=0.000	p=0.000	p=0.000	p=0.000
Weak-identification test						
Kleigbergen Paap rk Wald F statistic	25.214	13.184	11.172	23.669	12.068	11.187
Endogeneity test	p = 0.464	p=0.853	p = 0.293	p=0.417	p=0.824	p=0.239
Other Controls	Y	Y	Y	Y	Y	Y
District Fixed Effects	Y	Y	Y	Y	Y	Y
Time Fixed Effects	Y	Y	Y	Y	Y	Y
District*Income Fixed effects				Y	Y	Y
Time*Income Fixed effects				Y	Y	Y
N	29411	32824	18650	29411	32824	18650

* p<0.10, ** p<0.05, *** p<0.01. Panel A and B provide the DID and DDD results respectively. Estimation is using LPM. The sample in (1) is restricted to children between the ages 5 to 9 years; in (2) is restricted to children between the ages 10 to 14 years; and in (3) is restricted to children in the ages 15 to 17 years. Dependent variable is school enrollment of a child in a household in a district at a particular point in time. Additional controls included in each specification - gender dummy = 1 for a boy and 0 for a girl, RSBY = 1 if the district was exposed to RSBY, treatment interactions with gender dummy, Low Income dummy =1 HH does not belong to top 30% and 0 otherwise (for DDD), HH size, parental education characteristics, indicators for religion of HH, indicators for caste of HH, dummy for urban areas, school facilities and scholarships offered, district and time fixed effects, district by income fixed effects (for DDD) and time by income fixed effects (for DDD). HH size is instrumented by the gender of the first child. Standard errors reported are clustered standard errors.

Table 14. Sensitivity analysis: Impact of RSBY on school expenditure and child school enrollment - Rural vs urban

	Panel A. DID			Panel B. DDD		
Panel I - School expenditure	**Baseline**	**Urban**	**Rural**	**Baseline**	**Urban**	**Rural**
RSBY*Post	0.005***	0.007**	0.004***	-0.003*	-0.001	-0.004
	(0.001)	(0.003)	(0.001)	(0.002)	(0.004)	(0.003)
Low Income (=1 for bottom 70%)				-0.047**	0.000	0.001
				(0.024)	(0.006)	(0.003)
RSBY*Post*Low Income				0.007***	0.002	0.008**
				(0.001)	(0.005)	(0.003)
Other Controls	Y	Y	Y	Y	Y	Y
District Fixed Effects	Y	Y	Y	Y	Y	Y
Time Fixed Effects	Y	Y	Y	Y	Y	Y
District*Income Fixed Effects				Y	Y	Y
Time*Income Fixed Effects				Y	Y	Y
N	47421	11219	28897	47421	11205	28897
Panel II. School enrollment						
RSBY*Post	0.027***	-0.005	0.040***	-0.023	-0.043***	-0.027
	(0.006)	(0.011)	(0.007)	(0.017)	(0.014)	(0.018)
Boy	0.060***	0.023***	0.072***	0.055***	0.022***	0.068***
	(0.004)	(0.006)	(0.006)	(0.004)	(0.005)	(0.006)
RSBY*Post*Boy	-0.019***	-0.005	-0.024***			
	(0.005)	(0.008)	(0.006)			
Low Income (=1 for bottom 70%)				-0.042	-0.176	-0.008
				(0.023)	(0.164)	(0.105)
RSBY*Post*Low Income				0.046**	0.043**	0.066***
				(0.020)	(0.020)	(0.024)
RSBY*Post*Low Income*Boy				-0.009***	-0.007	-0.015**
				(0.001)	(0.012)	(0.007)
Underidentification test	p=0.000	p=0.000	p=0.000	p=0.000	p=0.000	p=0.000
Weak-identification test						
Kleigbergen Paap rk Wald F statistic	44.022	30.525	27.462	42.46	20.892	27.781
Endogeneity test	p = 0.570	p=0.402	p=0.566	p=0.414	p=0.598	p = 0.516
Other Controls	Y	Y	Y	Y	Y	Y
District Fixed Effects	Y	Y	Y	Y	Y	Y
Time Fixed Effects	Y	Y	Y	Y	Y	Y
District*Income Fixed Effects				Y	Y	Y
Time*Income Fixed Effects				Y	Y	Y
N	83221	22760	60461	83221	22760	60461

* p<0.10, ** p<0.05, *** p<0.01. Panel A and B provide the DID and DDD results respectively. Panel I provides the results estimated using IV approach. Dependent variable is budget share of household's school expenditure. The sample is restricted to HH with children and where age of the head is between 18 to 90 years. Panel II provides the results estimated using a LPM. Dependent variable is school enrollment of a child in a household. Individual sample is restricted to children above the age of 5 and below the age of 18. Additional controls included in Panel I include RSBY = 1 if the district was exposed to RSBY & 0 otherwise, dummy for Low Income =1 if HH does not belong to top 30% and 0 otherwise (for DDD), HH size (instrumented by gender of the first child), highest education degrees of male and female members, indicators for religion of HH, indicators for caste of HH, dummy for urban areas, number of married men in the HH, number of married women in the HH, proportion of children, teens and adults, indicator for if HOH is married, dummy for if the HH has a bank account, dummy for if the HH has a farmer credit card, district fixed effects, time fixed effects, district by income fixed effects (for DDD), time by income fixed effects (for DDD). Controls in Panel II specification include a gender dummy = 1 for a boy and 0 for a girl, RSBY, dummy for Low Income, HH size, parental education characteristics, indicators for religion of HH, indicators for caste of HH, dummy for urban areas, school facilities and scholarships offered, district and time fixed effects, district by income fixed effects (for DDD), time by income fixed effects (for DDD). HH size is instrumented by the gender of the first child. Standard errors reported are clustered standard errors.

Table 15. Sensitivity analysis: Impact of RSBY on school expenditure and child school enrollment - Variation by castes

	Panel A. DID						Panel B. DDD					
Panel I - School expenditure	**Baseline**	**General**	**OBC**	**SC**	**ST**	**Other**	**Baseline**	**General**	**OBC**	**SC**	**ST**	**Other**
RSBY*Post	0.005***	0.009***	0.004*	0.000	0.004	0.003	-0.003*	-0.009	-0.002	0.001	0.000	-0.005
	(0.001)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.002)	(0.008)	(0.003)	(0.003)	(0.006)	(0.004)
Low Income (=1 for bottom 70%)							-0.047**	-0.006***	-0.007**	-0.004	-0.008	-0.005
							(0.024)	(0.001)	(0.002)	(0.004)	(0.008)	(0.004)
RSBY*Post*Low Income							0.007***	0.019***	0.004	0.002	0.006	0.006*
							(0.001)	(0.006)	(0.003)	(0.003)	(0.006)	(0.003)
Other Controls	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
District Fixed Effects	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Time Fixed Effects	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
District*Income Fixed Effects							Y	Y	Y	Y	Y	Y
Time*Income Fixed Effects							Y	Y	Y	Y	Y	Y
N	47421	2206	15091	14565	7025	8534	47421	2206	15085	14565	7025	8534
Panel II. School enrollment												
RSBY*Post	0.027***	0.024	0.062***	0.046***	0.054	-0.015	-0.023	0.011	-0.024	-0.024	-0.134***	0.014
	(0.006)	(0.022)	(0.012)	(0.013)	(0.044)	(0.025)	(0.017)	(0.022)	(0.034)	(0.038)	(0.035)	(0.042)
Boy	0.060***	0.029***	0.077***	0.051***	0.041***	0.061***	0.055***	0.026***	0.064***	0.053***	0.032***	0.060***
	(0.004)	(0.010)	(0.009)	(0.008)	(0.014)	(0.009)	(0.004)	(0.010)	(0.009)	(0.007)	(0.010)	(0.008)
RSBY*Post*Boy	-0.019***	-0.011	-0.070***	0.008	0.001	0.001						
	(0.005)	(0.014)	(0.008)	(0.008)	(0.014)	(0.018)						
Low Income (=1 for bottom 70%)							-0.042	-0.472*	-0.014	-0.181*	-0.111	0.237
							(0.023)	(0.250)	(0.167)	(0.101)	(0.186)	(0.179)
RSBY*Post*Low Income							0.046**	0.034	0.072	0.070*	0.198***	-0.046
							(0.020)	(0.023)	(0.047)	(0.038)	(0.052)	(0.042)
RSBY*Post*Low Income*Boy							-0.009***	-0.007	-0.061***	0.004	-0.021*	-0.008***
							(0.001)	(0.020)	(0.011)	(0.009)	(0.013)	(0.001)
Underidentification test	p=0.000	p=0.000	p=0.000	p = 0.041	p = 0.040	p=0.000	p=0.000	p=0.000	p=0.000	p=0.056	p=0.003	p=0.000
Weak-identification test												
Kleigbergen Paap rk Wald F statistic	44.022	20.444	12.667	4.131	4.111	17.656	42.46	18.793	11.318	5.717	8.125	25.42
Endogeneity test	p = 0.570	p = 0.447	p = 0.530	p = 0.740	p = 0.605	p = 0.024	p=0.414	p = 0.433	p = 0.563	p = 0.849	p = 0.415	p = 0.010
Other Controls	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
District Fixed Effects	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Time Fixed Effects	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
District*Income Fixed Effects							Y	Y	Y	Y	Y	Y
Time*Income Fixed Effects							Y	Y	Y	Y	Y	Y
N	83221	4486	25892	25189	12291	15363	83221	4486	25892	25189	12291	15363

* p<0.10, ** p<0.05, *** p<0.01. Panel A and B provide the DID and DDD results respectively. Panel I provides the results estimated using IV approach. Dependent variable is budget share of household's school expenditure. The sample is restricted to HH with children and where age of the head is between 18 to 90 years. Panel II provides the results estimated using a LPM. Dependent variable is school enrollment of a child in a household. Individual sample is restricted to children above the age of 5 and below the age of 18. Additional controls included in Panel I specification include RSBY = 1 if the district was exposed to RSBY & 0 otherwise, dummy for Low Income =1 if HH does not belong to top 30% and 0 otherwise (for DDD), HH size (instrumented by gender of the first child), highest education degrees of male and female members, indicators for religion of HH, dummy for urban areas, number of married men in the HH, number of married women in the HH, proportion of children, teens and adults, indicator for if HOH is married, dummy for if the HH has a bank account, dummy for if the HH has a farmer credit card, district fixed effects, time fixed effects, district by income fixed effects (for DDD), time by income fixed effects (for DDD). Controls in Panel II include a gender dummy = 1 for a boy and 0 for a girl, RSBY, Low Income, HH size, parental education characteristics, indicators for religion of HH, dummy for urban areas, school facilities and scholarships offered, district and time fixed effects, district by income fixed effects (for DDD), time by income fixed effects (for DDD). HH size is instrumented by the gender of the first child. Standard errors reported are clustered standard errors.