I Z A Institute of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 17083

The Effects of Childhood Immunization Program on Health and Education: Micro-Evidence from India

Santosh Kumar

JUNE 2024

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 17083

The Effects of Childhood Immunization Program on Health and Education: Micro-Evidence from India

Santosh Kumar
University of Notre Dame, JPAL and IZA

JUNE 2024

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

The Effects of Childhood Immunization Program on Health and Education: Micro-Evidence from India*

Exploiting cohort and spatial variation in the exposure to the "Universal Immunization Program", I estimate the program's effects on child mortality and educational attainment in India. Results show that exposure to the program reduced infant mortality by 0.4 percentage points and under-five child mortality by 0.5 percentage points. While the program clearly reduced mortality, it had mixed effects on children's educational outcomes due to changes in the composition of children in the population. I find it had a negative impact on primary school completion, but a positive impact on secondary school completion. The negative effect at low levels of schooling may be due to lower average health among marginal surviving children or a quantity-quality trade-off where the unanticipated survival of children induces families to under-invest in each child. The greater propensity to complete secondary school on the other hand may be due to improved health among children who are farther away from the margin of survival.

JEL Classification: I1, I2, J18, O15, O22
Keywords: immunization, health, schooling, India

Corresponding author:
Santosh Kumar
Keough School of Global Affairs
University of Notre Dame
Notre Dame
IN, 46556
USA
E-mail: skumar23@nd.edu

* I thank Abhijit Banerjee, Aimee Chin, Chinhui Juhn, Elaine Liu, Rohini Pande, Farhan Majid, Anil Kumar, Carlos Jijon, Vikas Gawai, Andres Ramasco, Matt Hauenstein, and participants at several Economic Conferences and seminars for helpful comments and discussions. Any error remains mine.

1 Introduction

Investment in early childhood health has significant human capital returns in the form of higher education, better cognition, higher productivity, and better labor market outcomes (Alderman et al., 2001; Almond & Currie, 2011; Almond, Currie & Duque, 2018; Currie & Vogl, 2013). Childhood immunization has contributed immensely to the overall improvement of health, particularly from infectious diseases and associated mortality in developing countries. Childhood immunization programs are critical to public health and are the most cost-effective health interventions for reducing the disease burden among children globally. Immunization program further contributes to narrowing the health gap across socioeconomic groups. Therefore, childhood immunization has become a critical component of public health policies in many countries, yet immunization coverage is not universal (23 million children under the age of one year did not receive basic vaccines (WHO, 2021).[1]

In addition to the health benefits, it is of policy interest to understand the long-run effects of improvement in child health on educational outcomes. Although school enrollment and access to primary schools have grown considerably in developing countries, 263 million children, adolescents, and youth were out of school in 2019 (UNESCO, 2017). Most children never reach secondary school, and academic achievement as measured by standardized tests is dismally low (Glewwe & Kremer, 2006; Kremer, Brannen & Glennerster, 2013). The developing world is facing a learning crisis as over half of the 617 million in-school children in the primary and lower secondary age group worldwide lack the required minimum proficiencies (World Bank, 2018; Clarke, 2022). In a setting beset with high disease burden and poor learning outcomes among children, an important inquiry is to understand the extent to which poor child health is responsible for learning deficits in developing countries.

The evidence of immunization programs on education and other measures of human capital is limited and mixed. On the one hand, immunization programs may spur human capital formation as immunized children are more likely to attend schools, have a better attention span, and have better cognitive development (Bloom, Canning & Shenoy, 2012; Driessen et al., 2015; Nandi et al., 2020). On the other hand, immunization programs may have unin-

[1] For example, in 2019 about 65% of children under the age of 2 years were fully vaccinated in India (International Institute for Population Sciences, 2017)

tended consequences as a reduction in child mortality is likely to change the composition and size of the surviving pool of children. The unanticipated increase in the number of children may strain the educational infrastructures and thus reduce the quality of education. This quantity-quality trade-off channel does not rule out the possibility of deleterious effects on human capital formation, especially in countries with weak state capacity and sub-optimal educational infrastructures to deliver quality education.

This study evaluates the impacts of a large-scale government-sponsored childhood immunization program in India called the "Universal Immunization Program" (hereafter, UIP) on child mortality and educational attainment. In 1985-86 the Government of India launched UIP in 31 districts. Each year additional districts were phased into the program and by 1990 all 443 districts of India were covered by UIP. The program delivered free immunization shots for children under one year of age to protect them from six Vaccine-Preventable Diseases (hereafter, VPDs).[2] I exploit the staggered roll-out of the UIP program and use cohort and district variation in the program exposure to estimate UIP's effects on mortality and education. The empirical methodology is akin to the difference-in-difference (DID) method and plausibly identifies the causal impacts of the UIP program.

UIP has features that facilitate evaluation. On the one hand, there is district variation in UIP exposure: UIP was implemented gradually across districts in India, with the timing apparently determined by fixed district characteristics. On the other hand, there is cohort variation in UIP exposure: only children who were twelve months old or younger at the time the program began would have been eligible to receive immunizations. This enables me to use a difference-in-differences-type estimation strategy to identify the effect of UIP. The identifying assumption is that without UIP, the cohort difference in mortality and educational outcomes would have been the same between the districts that implemented UIP sooner and the districts that implemented UIP later. I apply this identification strategy using the "Reproductive and Child Health Survey" (hereafter, RCH), a large nationally representative individual-level data set.

The main finding of this paper is that UIP reduced infant mortality by 0.4 percentage points and under-five mortality by 0.5 percentage points. The effects on mortality outcomes

[2]The six VPDs are Diphtheria, Pertussis, Tetanus, Poliomyelitis, Measles, and Tuberculosis.

are substantial given that the infant mortality in India was 9.7% and the under-five mortality estimate was 15% before the launch of the program. Indeed they account for approximately one-fifth of the decline in infant and under-five child mortality rates between 1985-1990. I also employ the recent econometric method Callaway and Sant' Anna (2020) to show the robustness of my main findings (figure 3). Furthermore, there is no differential effect by gender but effects are more pronounced in rural areas and for poor and low-caste children. I verify that these results are not due to differential trends in child mortality between early-UIP districts and later-UIP districts by doing placebo tests using older cohorts and using a health outcome unrelated to the immunizations.

Next, I examine the effects of UIP on the educational outcomes of surviving children and find mixed results. The program had a negative impact on primary school completion, but a positive impact on secondary school completion. The results on education outcomes can be explained in terms of changes in the composition of the surviving children due to the immunization program. The negative effect on education may be due to the lower quality of the "marginal child" is similar to the argument made by Donohue and Levitt (2001) and Gruber, Levine, and Staiger (1999); UIP induced some children to survive who otherwise would have died, and these children may be less healthy. The negative results are also consistent with the quantity-quality trade-off where an unanticipated increase in household size due to the immunization program induces the households to under-invest in each child. Furthermore, the quality of education represents an additional pathway explaining the nuanced results in this paper. The UIP-led decline in child mortality increased the number of surviving children. This, in turn, could have led to overcrowded classrooms, exerting a negative impact on education quality. The deleterious consequences of classroom overcrowding and a deteriorating student-teacher ratio in developing countries are widespread, with prior research indicating that a substantial surge in primary school enrollments exacerbated educational quality issues (Pritchett 2013, World Bank 2018).

On the other hand, the result that UIP increased the education of some children is similar to the evidence on the positive effects of disease eradication or improved morbidity on human capital outcomes (Miguel and Kremer, 2004; Cutler et al., 2010; Lucas, 2010; Bleakley, 2007; Ozier, 2018; Hamory et al., 2021). The greater propensity to complete secondary school may be due to improved health among those children who are not at the margin of survival.

This paper contributes to the existing literature in several ways. First, I am not aware of any previous studies that rigorously quantify the effects of immunization programs on the health and education outcomes of children.[3] On the one hand, there have been process evaluations that describe the implementation of UIP and vaccination coverage. On the other hand, there have been medical evaluations that examine the effects of immunizations on health outcomes but these studies are conducted in laboratory-like settings rather than in real-world developing-country contexts. There is widespread skepticism about the public health service delivery system in developing countries. Chaudhury et al. (2006) find, for instance, that 39% of doctors and 31% of other health care workers were absent from work in nationally representative surveys of primary health centers in Bangladesh, Ecuador, India, Indonesia, Peru, and Uganda. In these environments, can a universal immunization program successfully reduce child mortality and what are the consequences for children's educational outcomes remain an open inquiry.

Second, this paper adds to the literature on the effects of child health on schooling. There have been several studies examining the effects of improved health on human capital (Miguel and Kremer, 2004; Lucas, 2010; Bleakley, 2007; Ozier, 2018). However, unlike the disease eradication program, I used variation in health provided by a large immunization program. Immunization programs provide a new and different source of variation in child health. In particular, whereas health interventions used by other studies primarily reduce the morbidity of children, immunization programs reduce both the mortality and morbidity of children. Since immunization programs operate on different margins, the consequences for education and human capital formation could be different from those other health interventions such as malaria eradication, deworming, etc.

This study differs from Nandi et al. (2020) in several important ways. Nandi et al. (2020) examined the long-term impacts of UIP on schooling and found a significant association between exposure to UIP and schooling. Compared to Nandi et al. (2020), this study uses different empirical methods and analyzes mortality outcomes in addition to schooling out-

[3] Nandi et al. (2020) compared the pre-and post-UIP cohort and found that years of schooling increased by 0.25 years in the fixed effect models. My study uses different empirical methods and analyzes mortality outcomes in addition to schooling outcomes.

comes. I use an advanced econometric technique that is at the research frontier of staggered DiD–when there is variation in treatment timing. Recent studies cast doubt on the validity and robustness of two way fixed effect (TWFE) DiD estimator when the rollout of the program is staggered over time (i.e. there are multiple treatment periods) (Sun and Abraham, 2020; Callaway and Sant'Anna, 2020). Since UIP rollout was staggered over time, this study provides estimates robust to parallel trend violations and heterogeneous treatment effects. Furthermore, I exploit within-district cohort variation in the UIP exposure and employ the DiD method that allows me to estimate the causal impacts of the UIP on the outcomes, while Nandi et al. (2020) compared pre-UIP to the post-UIP cohort to estimate the impacts on schooling. Nandi et al. (2020) use an adult sample who had completed schooling, while this study includes adolescents who are still in school (the mean age is 13 years)–an age groups more amenable to policy intervention. Finally, I decompose schooling effects by years of schooling (at each grade level), while Nandi et al. (2020) do not investigate schooling effects by grades.

The rest of the paper is structured as follows: in Section 2 I discuss the related literature and provide an overview of UIP. Section 3 presents the empirical framework and Section 4 describes the data. Section 5 presents the results on mortality outcomes and Section 6 presents the results on educational outcomes. Finally, Section 7 concludes.

2 Background

2.1 The Universal Immunization Program

Approximately 3 million children died each year of VPDs with a disproportionate number of these children residing in low- and middle-income countries (Kane and Lasher, 2002). Of the estimated 5.3 million child deaths in 2018, approximately 700,000 die each year of VPDs. Vaccines remain one of the most cost-effective public health initiatives, yet the cover against VPDs remains far from complete; recent estimates suggest that approximately 23 million children are not completely immunized with almost 98 percent of them residing in developing countries (WHO, 2021). The global Millennium Development Goal of reducing child mortality by two-thirds by 2015 had some success but the Sustainable Development Goal of reducing under-5 mortality to at least as low as 25 per 1,000 live births can only be met if all children are universally immunized and there is a significant reduction in VPDs.

In India, immunization of children against VPDs has been a central the goal of the health care system from the 1970s. The Expanded Program on Immunization (EPI) was initiated in 1978 to make six childhood vaccines (BCG, DPT, TT, DT, Polio, and typhoid) available to all eligible children. The main objective of EPI was to reduce mortality and morbidity by controlling six target diseases- Tuberculosis, Diphtheria, Tetanus, Pertussis, Polio and Typhoid. EPI failed to achieve the objective of immunizing children; because the program was limited primarily to major hospitals in urban areas and coverage levels were very low. In 1985, the Government of India made childhood immunization a Technology Mission and launched UIP with much dynamism to attain the goal of achieving 85 percent coverage for tuberculosis, diphtheria, tetanus, pertussis, polio and measles for all children by 1990.

Under UIP, each child had to be vaccinated before he or she turned one year of age with three doses of the DPT vaccine, three doses of the polio vaccine, and one dose of each of the measles and BCG vaccines. Table 1 in the appendix lists some symptoms associated with the diseases that these shots protect against. The symptoms range from mild to severe, with serious sickness and death more likely among infants (whose immune systems are not yet mature) and poor children (whose immune systems are weakened due to malnutrition). It is worth noting that immunization protects individuals not only from illness per se but also from the long-term effects of that illness on their physical, emotional, and cognitive development (Bloom, Canning, & Weston, 2005; Bloom, 2011). Additionally, these diseases are communicable, so there are significant positive externalities from being vaccinated. That is, the vaccines reduce the risk of disease not only for the vaccinated children but also for people around them by reducing the transmission rate of the diseases.

There were not sufficient resources to implement the program all over the country at the same time. Thus, UIP had a phased roll-out, beginning with 31 districts in 1985-86 and covering all districts by 1990. The program was implemented through the existing network of primary health care infrastructure which consists of a referral center called "community health center" for every 80 to 120 thousand people, a primary health center for 20 to 30 thousand people, and a sub-center for every 3 to 5 thousand people. The program made provision for additional inputs in the form of additional staff, vaccines, and equipment for storage and transportation of vaccines such as walk-in-coolers, refrigerators and vaccine carriers.

Below I take advantage of the staggered implementation of UIP across districts to help identify UIP's effect, therefore it is essential to understand what determined the timing. Toward this end, I had numerous conversations with officials in the UIP division of the Ministry of Health and Family Welfare. The timing was not completely random. It seems that the capacity of the district to achieve the immunization coverage rates targeted by UIP and maintaining this level in subsequent years was a major factor in the selection of the district. In addition, infrastructure and other health facilities to deliver the UIP services were also taken into account while selecting the districts. In other words, the selection of districts was based on fixed characteristics of the districts. For example, early-adopting districts may have more primary health centers, more nurses, or better healthcare infrastructure. The selection on fixed district characteristics does not cause problems for the interpretation of my estimated treatment effects because they rely on within-district variation in exposure to UIP only; that is, I always control for district fixed effects. A more serious problem would be if the timing of implementation depended on underlying district-specific trends in the outcome variables. It must be emphasized that UIP officials never indicated that district trends in mortality or education were part of the criteria for earlier implementation. However, to address this potential concern, I perform control experiments using older cohorts who are not exposed to UIP and using a health outcome unrelated to the immunizations; I discuss these control experiments in detail in section 5.

UIP is one of the largest vaccination programs in the world in terms of quantities of vaccines used, number of intended beneficiaries, number of immunization sessions organized, the geographical spread, and the diversity of areas covered. Surprisingly, there have not been many studies estimating the impact of UIP's effect on mortality, let alone its potential effects on education. Previous evaluations of UIP were mostly sanctioned by the Ministry of Health and Family Welfare and international donor agencies like WHO, UNICEF and were basically process evaluations that look at the coverage of vaccines.[4] They show that UIP was able to substantially increase the coverage of immunization shots (Figure 1 in Appendix). Vaccine

[4]Gupta and Murali (1989); Sathyamala (1989); Annual Report (1987-88), MoHFW; UNICEF, 2002).

coverage by antigen shows a substantial increase during the UIP period. The vaccine coverage increased from a low 30-40% at the start of the program to approximately 80-100% by 1990-91.

The extent to which UIP reduced child mortality and its effects on the schooling of surviving children remains an open question. Answering these questions is of great interest to India. There is widespread debate about the efficient implementation of public health programs in India. Many claim that the public health service delivery system in India is inefficient and that government-sponsored programs exist only on paper and that real take-off of public health programs is either doubtful or slow. Moreover, it should have relevance for policymakers and public health activists outside of India. Many countries have mass immunization programs or are considering adopting them. Immunization programs compete for limited funding for public health initiatives and other welfare programs, with budget constraints especially tight in poor developing countries. Do immunization programs help the intended beneficiaries? I present my strategy for evaluating UIP after briefly reviewing the related literature below.

2.2 Related Literature

This paper lies at the intersection of two strands of the literature which I discuss briefly below. One on the effectiveness of the immunization program on child health and the other on the effects of early child health on human capital accumulation and how a reduction in disease burden brought out by lower child mortality may affect education in different ways.

In the public health discipline, there have been several efficacy studies or clinical trials of vaccines. These small-scale clinical trials have unequivocally established that the DPT vaccine effectively protects against diphtheria, pertussis, and tetanus; the BCG vaccine against typhoid and the measles and polio vaccines against measles and polio, respectively (Levine et al., 2005; United Nations, 2006). There have also been a few epidemiological studies in developing countries that examine the impact of specific vaccines on child mortality (Breiman et al., 2004 for Matlab, Bangaladesh; Koeing et al., 1990 for Senegal). These studies reconfirm the laboratory evidence and find the decreased risk of death for vaccinated children. The sample sizes tend to be small in these studies, unfortunately. *Finally, numerous studies examine* *the cost-effectiveness and cost-benefit of vaccines* (*Navas, 2005; Ekwueme, 2000; see Bloom, Canning, and Weston, 2005 for a review*). These studies look at outcomes like averted illnesses, hospitalizations, and deaths, disability-adjusted life years (DALYs) gains, and medical costs. These studies suggest immunization is a highly cost-effective intervention. This medical research underlies the public health policy of many countries to require vaccinations for all children.

This paper differs from medical studies in several respects. First, it is evaluating the effect of a program that provided vaccinations, not the effect of the vaccinations as the medical studies have done. Although we know scientifically that vaccinations reduce mortality, we do not know whether a scaled-up mass immunization program can be effective in reducing mortality in a resource-constrained country such as India. The success of the program depends not only on the efficacy of the vaccinations but also on the public health delivery system. Second, given the low baseline vaccination coverage and average health status in poor developing countries, it may well be that the marginal impact of the vaccinations may be greater than in developed countries where the medical studies were done. Third, the medical literature has ignored non-health outcomes such as education.

There is a large body of literature in economics that shows the causal impacts of health and education (Almond, Currie, & Duque, 2018). Researchers have used randomized experiments and natural experiments to identify the causal effect of health on education. In an early work, Dow, Philipson, and Sala-i-Martin (1999) provide evidence that the decrease in mortality risks resulting from the World Health Organization's Expanded Programme on Immunization led to an increase in parental health investments and highlight the importance of spillover effects on mortality and morbidity of other causes of deaths.

In randomized experiments, the researcher randomly assigns similar units to different health treatments (a treatment group that receives medical treatment and a control group that doesn't), generating an exogenous source of variation in health. Randomized experiments examining the effect of child health on education include the following. Miguel and Kremer (2004) find that providing children with deworming medication significantly reduced serious worm infections and increased school attendance. Bobonis, Miguel, and Sharma (2006) find that an iron supplementation program significantly reduced anemia and school absenteeism. In contrast, Krämer, Kumar & Vollmer (2021) find no effects of a reduction in anemia on attendance, cognition, and test scores among second-grade children in India.

Researchers have also used natural experiments to identify the effect of child health on educational outcomes. Along these lines, Bleakley (2007) exploits region-by-time variation in exposure to the hookworm eradication program sponsored by the Rockefeller Sanitary Commission in the 1910s in the United States to identify the effects of reducing hookworm infections on educational outcomes. He finds that regions that experienced greater reductions in hookworm infections had larger increases in school attendance and literacy. Other studies have also used a difference-in-differences strategy to estimate the effect of malaria on human capital accumulation (Bleakley, 2010 for the U.S., Brazil, Colombia, and Mexico; Cutler et al., 2010 for India; Lucas, 2010 for Sri Lanka and Paraguay).

This paper has an empirical design similar to studies in the latter group–takes advantage of a natural experiment to identify the effect of health on education. As described in the next section, I use district-by-cohort variation in exposure to UIP to obtain estimates of the effect of health on education. It is one of only a handful of studies that address the issue of endogeneity in health when estimating the effect of health on education. Furthermore, I use a new and different source of variation in child health that has rarely been used in prior studies. In particular, a few studies have looked at the impact of large-scale childhood immunization programs (Nandi et al., 2020). Childhood immunization prevents both mortality as well as morbidity from infections that disproportionately affect children. The most significant impact of vaccines has been to prevent morbidity and mortality from serious infections that disproportionately affect children. In contrast, the aforementioned studies that estimate the causal effect of health on education use health interventions–deworming treatments, malaria treatments, and school meals for school-aged children that operate primarily on the morbidity margin.

3 Data

My empirical analysis uses data from two sources: individual-level data from the Reproductive and Child Health (RCH) Survey and administrative data about UIP from the Ministry of Health and Family Welfare, Government of India.

The RCH survey is a large, nationally representative survey. Due to the timing of UIP, it is appropriate to use the second wave of the RCH survey, which was conducted during 2002-2004. First, I use the "fertility file" to construct the sample for my child mortality analysis. [5] Fertility history is collected for one woman who is aged 15 to 44 from each surveyed household. The fertility history includes information on all children ever born to a woman even if the child has died by the time of the survey. This enables me to collect for each child born to a woman in the fertility file the year of birth, whether alive at the time of the survey, and the year of death for children who died before the survey year. Second, I use the "household file" to construct the sample for analyzing the educational outcomes of surviving children. Both the fertility and household files contain information on such control variables as district, rural/urban, child sex, age, and birth order, and household social group, religion, and socio-economic conditions.

The Ministry of Health and Family Welfare provided administrative information about UIP. First, I held several rounds of discussions with several UIP officials to find out the details of how UIP was implemented. It was these conversations that led me to believe that the timing of UIP could be considered exogenous conditional on district fixed effects, leading me to adopt the difference-in-differences strategy. Second, I obtained from them a list of new districts that implemented UIP each year, from year 1 (1985-86) to year 5 when all districts were covered by the UIP (1989-90).

I mapped the year of UIP implementation from the district-level administrative data back to the individual-level RCH survey data using the district codes. One complication was that the number of districts increased from 443 to 593 during the UIP period (1985-1990) and the survey period (2002-04). Either an existing district was split into two or more new districts or a new district was formed by taking areas from two or more districts. I successfully match 563 districts by looking at district census handbooks, district websites, and other government sources (a success rate of 95 percent).

The main outcome variables for my mortality analysis are the probability of dying within

[5] The RCH survey is a repeated cross-section (at the district level) and has 4 waves starting from 1998-99, 2002-04, 2007-08, and 2012-13 (only in a few selected states). The second wave of RCH is most suitable for our analysis as it contains retrospective mortality history of children going back to UIP years. The later rounds of the RCH survey do not have enough observations on child mortality for children born in UIP years. The first wave of the RCH survey is also not suitable because many UIP children would still be in school and would not have completed primary or secondary schooling.

the first twelve months-Pr(Infant Mortality)- and the probability of dying within the first five years–Pr(Under-five Mortality). Both infant mortality and under-five mortality are common health indicators used by governments and international agencies for tracking improvement in population health and disease burden. These indicators are also SDG goals. Under-five mortality is an indicator of the cumulative exposure to mortality risk during the most vulnerable years of childhood; it includes infant mortality as well as mortality from age 1 to 5.

For the analysis of the educational outcomes of surviving children, the main outcome variables are Pr(Literate), Pr(Primary School Completion), Pr(Middle School Completion), Pr(Secondary School Completion) and Years of Schooling. All the education outcomes variables are dichotomous variables except Years of Schooling and are defined as follows. Pr(Primary School Completion) is defined as Pr(Years of Schooling $\geq$5), Pr(Middle School Completion) is defined as Pr(Years of Schooling $\geq$8), Pr(Secondary School Completion) is defined as Pr(Years of Schooling $\geq$10). The education outcomes variables are conditional on being literate.[6]

Table 1 shows the descriptive statistics of the variables used in the mortality and education analyses. The paper uses children born between 1983 and 1992 for the mortality outcomes and children born between 1983 and 1997 for the educational outcomes. The number of observations for child mortality analysis is 297,385 and for education analysis, there are 898,789 observations.[7] In the child mortality sample, 69 percent of the children live in rural areas, and 47 percent belongs to a poor household. The majority of the children are Hindu (76 percent) and disadvantaged minority groups ST and SC forms 33 percent of the sample. The mean mother's age is 37.1 years and only 39 percent of the mothers are literate. The mean mother's age is higher than the mother's age of the average child because the survey was done in 2002-04 and the paper uses the children born between 1983 and 1992. The demographic and family characteristics of the sample used in the education analysis are similar to the sample used in the child mortality analysis.

In the child mortality sample, the mean infant mortality rate is 9% and the mean under-

[6]The survey asks Years of Schooling questions only to literate individuals.

[7]Different samples are used for the child mortality analysis and education analysis because as explained earlier in this section, the data for the mortality analysis are from the "fertility file" of the RCH and the data for the education analysis are from the "household file". Only one woman aged 15-44 from each household answered the supplemental fertility questions, hence the smaller sample size.

five mortality rate is 11%. In the education sample, 82 percent of the children are literate, and conditional on being literate the average years of schooling is 6.2 years. Conditional on being literate, about 65 percent children have completed primary schooling, 36 percent have completed middle school and only 18 percent of children have completed secondary school. The mean age of the children is 13.71 years; some of the children in the sample are still in school.

4 Empirical Framework

The objective of this study is to estimate the causal impact of a mass immunization program on child mortality and education outcomes in India. Ideally, to identify the effect, we would conduct a randomized experiment where some children are placed into a treatment group that receives immunizations and others are placed in a control group. We would follow these children over time and compare their mortality and educational outcomes. The control group describes the counterfactual of what the treatment group's outcomes would have been had the health intervention not occurred. This is a simple and convincing approach since, at the outset of the experiment, the children were similar across intervention arms.

In the absence of a randomized experiment, I rely on a natural experiment. I use the variation provided by India's implementation of UIP in the 1980s. In particular, I estimate the program effect by utilizing the following two sources of variation in exposure to UIP: variation across districts and variation across cohorts. First, variation across districts comes from the fact that districts got the program in different years. Figure 2 in the appendix shows the number of districts added to UIP each year. UIP was implemented in 48 districts (31 according to old district definitions) in the first year, 92 additional ones in the second year, and so on until all 563 districts (443 according to old district definitions) were covered in 1990.[8] Second, variation across cohorts comes from the fact that only children who are twelve months or younger when UIP was implemented would have been eligible to receive the shots. Table 2 in the appendix shows the schedule for the vaccines that UIP provided; the shots are administered on a strict schedule in the first year of a child's life for maximal efficacy.

[8]The number of districts increased from 443 to 593 between the UIP period (1985-1990) and the RCH survey year (2002-04). The data section has more details.

Children older than one year were not treated by UIP. Table 3 of the The appendix shows the birth cohorts that were eligible for UIP by the district's year of UIP inception. For example, a child born in 1985 would have been exposed to UIP if he lived in one of the 48 districts that implemented UIP first (in 1986), but not if he lived in a district that implemented UIP later.

4.1 Difference-in-Differences Strategy

Exploiting the within-district cross-cohort variation in the exposure to UIP, I employ the difference-in-differences approach uses to identify the effect of UIP on mortality and educational outcomes.

Consider the following equation:

$$Y_{idt} = \beta_0 + \beta_1 UIP_{idt} + \delta X_{idt} + e_{idt} \tag{1}$$

where Y_{idt} is the outcome variables for individual i, residing in district d, at time t. UIP_{idt} indicates exposure to UIP and X is a vector of individual and household characteristics (e.g., sex, caste, religion, age, birth order, mother's education, mother's age, rural/urban). e_{idt} is the error term.

I wish to estimate the effect of UIP so the parameter of interest in this equation is β_1. But β_1 in equation (1) may not be consistently estimated due to omitted variable bias. Districts may be different from each other on many unobserved dimensions that can affect the outcome variables. For example, UIP officials indicated that the timing of UIP implementation across districts were not random and instead apparently determined by some features of districts that could be considered constant over time such as health infrastructure. Similarly, each year of birth can also be systematically different in ways that affect the outcome variables. For example, there is progress in health and education over time, or some country-wide economic shocks in a particular a year may affect that year's newborns differently than the infants born in another year. I address these concerns by adding district fixed effects and year of birth fixed effects to equation (1):

$$Y_{idt} = \beta_0 + \beta_1 UIP_{idt} + \gamma_d + \phi_t + \delta X_{idt} + e_{idt} \tag{2}$$

where Y_{idt} is the outcome variables for individual i, residing in district d, at time t. UIP_{idt} indicates exposure to UIP. γ and ϕ are district and year of birth fixed effects, respectively. X is individual and household control variables. All standard errors reported in this paper are robust and clustered by the district. These adjustments allow for conditional heteroskedasticity and for conditional autocorrelation within districts (Bertrand et al., 2004)

The parameter β_1 in equation (2) can be interpreted as the causal effect of UIP under the assumption that the difference in outcomes between the younger and older cohorts would have been the same between earlier-implementing districts and later-implementing districts in the absence of UIP. [9] In other words, the parallel trend assumption should be satisfied. While it is not possible to directly test this assumption–it is a counterfactual–I assess its validity in a couple of ways. Below, I do this by: (1) estimating equation (2) using only older cohorts that have never been exposed to UIP but where I falsify their treatment status; and (2) estimating equation (2) using an outcome that is unlikely to be affected by the program.

The identifying assumption would also be violated if some other contemporaneous interventions had the same district-by-cohort variation as UIP and which also affects the outcome variables I examine. To the best of my knowledge, I am not aware of any child health or education intervention that can contaminate the identification of the effect of the UIP.

Furthermore, the parameter β_1 in equation (2) may underestimate the true effect of UIP for a couple of reasons. First, because the consumption of vaccines has a positive externality, the control group may benefit indirectly from UIP. That is, although the control group is not eligible for UIP vaccinations, they may benefit because the diseases spread more slowly when more people are vaccinated. Second, recall that the EPI program preceded the UIP program in India. The EPI had very low vaccination coverage rates and operated only out of major hospitals so it is unlikely to pose a significant problem. But it is quite likely that the control (i.e., older) cohorts in urban areas got vaccinated under EPI, which means the treatment group is partially treated already. This would cause me to underestimate the program's effect in urban areas. Third, inter-district movement of households across districts may confound

[9] If UIP exposure was a simple interaction between two binary variables, say being in an earlier-implementing district and being in a younger birth cohort, then β_1 would be a differences-in-differences estimate, i.e., the cohort difference in outcome in earlier-implementing states that are in excess of the cohort difference in later-implementing states. In fact, I use more variation in UIP exposure but the intuition is similar to the simple binary case and so I term my approach a difference-in-differences-type strategy.

the estimated effects, but it is not much of a concern here, because of limited inter-district migration in India (Munshi and Rosenzweig, 2006; Chatttopadhyay and Duflo, 2004).

4.2 Allowing for Heterogeneity in Program Effects

UIP may not have uniform impacts. The impact may differ based on child sex, socioeconomic status, rural/urban, caste, etc. For example, perhaps the rich would have immunized their children even in the absence of the program and it is the poor who would benefit more from the program. Furthermore, there has been some evidence of "elite capture" in public and poorly funded health schemes in developing countries and it is possible that the rich and elite class capture most of the benefits of the program. As another example, the program may have different impacts in rural areas from urban areas due to differences in the availability of health care infrastructure to deliver the services. Also, the program effect may vary by gender because of intrahousehold discrimination in the allocation of resources by gender. For example, Oster (2007) shows that girls in India are discriminated against in access to vaccination.

To test whether the program effect varies by sex, I estimate the following equation:

$$Y_{idt} = \beta_0 + \beta_1 UIP_{idt} + \beta_2 UIP_{idt} * Female_{idt} + \beta_3 Female_{idt} + \gamma_d + \phi_t + \delta X_{idt} + e_{idt} \quad (3)$$

where the omitted category is male. β_1 measures the average program effect for male children and β_2 captures the additional program effect for females.

Similarly, to examine whether there is a differential program impact by rural residence, I estimate the following equation:

$$Y_{idt} = \beta_0 + \beta_1 UIP_{idt} + \beta_2 UIP_{idt} * Rural_{idt} + \beta_3 Rural_{idt} + \gamma_d + \phi_t + \delta X_{idt} + e_{idt} \quad (4)$$

where the omitted category is urban residence. β_1 captures the program effect for children residing in urban areas and β_2 captures the additional program effect of residing in rural areas.

Next, I allow the program effect to vary by the socio-economic status (SES) of the household.[10] To capture the heterogeneity in treatment effect by household SES, I estimate the following equation:

$$\begin{aligned} Y_{idt} = \beta_0 + \beta_1 UIP_{idt} + \beta_2 UIP_{idt} * Low_{idt} + \beta_3 UIP_{idt} * Middle_{idt} \\ +\beta_4 Low_{idt} + \beta_5 Middle_{idt} + \gamma_d + \phi_t + \delta X_{idt} + e_{idt} \end{aligned} \tag{5}$$

where the omitted category is households with high SES. β_1 captures the program effect for children residing in a household with high SES, β_2 captures the differential program effect for children residing in household with low SES (relative to the high SES category) and β_3 captures the differential program effect for children residing in households with middle SES.

To estimate how the program effects differ by social group, I estimate the following equation:

$$\begin{aligned} Y_{idt} = \beta_0 + \beta_1 UIP_{idt} + \beta_2 UIP_{idt} * SC_{idt} + \beta_3 UIP_{idt} * ST_{idt} + \beta_4 UIP_{idt} * OBC_{idt} \\ + \beta_4 SC_{idt} + \beta_5 ST_{idt} + +\beta_6 OBC_{idt} + \gamma_d + \phi_t + \delta X_{idt} + e_{idt} \end{aligned} \tag{6}$$

where the omitted category is households from high caste groups. β_1 captures the program effect for children belonging to high caste groups, β_2 captures the differential program effect for children belonging to Schedule Castes (SC) (relative to the high castes), β_3 captures the differential program effect for children belonging to Schedule Tribes (ST) and β_4 captures the differential program effect for children belonging to Other Backward Castes (OBC). SCs and STs are socially disadvantaged groups in India. OBCs are not as poorly off as SCs, but also have faced discrimination historically.

[10]The survey asks whether the household owns the following consumer durables: radio, television set, refrigerator, bicycle, motorcycle, and car. Based on ownership of these consumer durables, the RCH survey categorizes the household into three different categories in terms of SES: Low, Middle, and High. Though this is not the perfect measure of household wealth status, this is the best we can do given the fact that the health survey does not collect direct information on the income and wealth of the households.

5 Effect of UIP on Child Mortality

5.1 Basic Results

First, I show a balance test comparing early- and late-UIP districts to check the potential selection bias in the results in Table 2. For example, if the program was first implemented in poorer regions, the marginal effect would be higher, on the other hand, if the rollout started first in more developed regions the estimated effect would be lower. I categorized early UIP districts as districts that received UIP in 1985 and 1986, while late UIP districts are the districts that received UIP after 1986. Results in Table 2 show that there are no systematic differences between early- and late-UIP districts based on the 1981 census districts' characteristics. The p-values in column 4 are mostly > 0.10, indicating statistically insignificant differences between early- and late0UIP districts

Table 3 reports the results of estimating equation (1) and equation (2) for infant mortality and for under-five mortality.[8] The main coefficient of interest is the coefficient for the variable "Exposed to UIP", which gives the average program effect. Column (1) and Column (4) show the program effect without the control variables. e Estimates from Column (1) and Column (3) suggest a significant negative impact of the program on Pr(Infant Mortality) and Pr(Under-Five Mortality). Results from columns (1) and (3) suggest that the program decreases the probability of infant mortality by 0.9 percentage points (pp) and the probability of under-five mortality by 1.2 pp. Columns (2) and (4) show the results with controls and columns (3) and (6) show the results with controls as well as district and year of birth fixed effects.

On comparing the naive estimates with difference-in-differences estimates in Column (3) and Column (6), it turns out that naive estimates grossly overestimate the program effect. Estimating equation (1) is not the correct approach because districts differ in their fixed characteristics and ignoring these fixed differences would bias the program effect. In particular, UIP officials stated that the timing of UIP was not random, and instead districts that had better healthcare infrastructure received it sooner. Given this statement, it is not surprising that

[8] I estimate these models using OLS, i.e., using the linear probability model with standard errors clustered at the district level. I also estimated these models using logit and find qualitatively similar results; these results are available upon request.

the estimates in columns (2) and (4) are so large-they encapsulate not only the true effect of UIP but also the effect of being in a district with better health care infrastructure (which not surprisingly has a large negative effect on mortality). Similarly, there could be cohort-specific characteristics and without taking into account of these characteristics, it is not possible to get the true effect of the program. For example, there is improvement in health conditions and care in India over time, so younger cohorts would have lower mortality even without UIP, and the estimates in columns (2) and (4) erroneously attribute these secular improvements over time to UIP.

The naive estimates in columns (2) and (4) highlight the dangers of giving causal interpretations to parameter estimates when the sources of variation are not plausibly exogenous, and motivate my difference-in-differences approach. The preferred estimates are in columns (3) and (6)-these are results from estimating equation (2), which includes district fixed effects and year of birth fixed effects. Results from columns (3) and (6) suggest that the program significantly reduces infant mortality and under-five mortality. UIP decreases the probability of infant mortality by 0.4 pp and the estimate is statistically significant at a 10 percent level of significance (column 3). The official estimate of infant mortality was 9.7 pp in 1985 and 8 pp in 1990. The 0.4 pp effect of UIP is 4.1% of the baseline infant mortality rate and a fifth of the decline in infant mortality between 1985-1990. Thus, over a short period, UIP caused a meaningfully sized reduction in infant mortality; it took India thirty-four years to bring down infant mortality from 14.6% in 1951 to 9.7% in 1985.

Column(6) reports the results for under-five mortality. Results show that the program has a negative and significant impact on under-five mortality. The program reduced under-five mortality by 0.5 pp. The under-five mortality was 15 pp in 1985 and reduced to 12.3 pp by 1990. The 0.5 pp effect of UIP is 3.3% of the baseline under-five mortality rate and almost a fifth of the decline in under-five mortality between 1985-1990. It should be noted that the program has a larger impact on under-five mortality compared to infant mortality. This is to be expected because the under-five mortality rate includes infant mortality as well as mortality of children aged 1 up to 5. Thus there is a tenth of a percentage point decline in the mortality of children aged 1 up to 5 (the fourth tenth of a percentage point is the decline in mortality of children up to age 1). It makes sense that the mortality declines are greatest for infants. Vaccinations under UIP begin at birth, and though maximal protection is not gained until all the doses are administered according to schedule, protection begins right away. This early protection makes a big difference for infant survival since infants do not have well-developed immune systems yet.

In all the regression models in Table 3, the signs of the control variables are as expected. Mother's age and mother's education have a negative and significant effect on infants and under-five mortality. Poor and disadvantaged minority children(ST and SC) are more likely to die. For Other Backward Caste and Hindu children, the estimates are positive and significant, meaning that children belonging to these categories have a higher probability of dying.

5.2 Event Study Figures

One may worry about the plausibility of the parallel trends assumption in our TWFE empirical design. One could worry that early-UIP districts might have a differential trend in mortality. I address this concern by estimating a fully dynamic version of equation (2). Figure 4 in the appendix shows the event study estimates of the TWFE model with indicators for time to treatment. The top figure shows the event-study version of the TWFE model. However, TWFE estimates do not provide consistent estimates if there is heterogeneity in treatment effects (Callaway and Sant' Anna 2020; Sun and Abraham 2021). I present the event study figures that are robust to treatment effect heterogeneity. Figure 3 shows the event study estimator proposed by Callaway and Sant'Anna (2020) and the bottom chart in Figure 4 shows the estimator proposed by Sun and Abraham (2021). The event study figures show that the DiD estimates in the Pre-UIP periods are statistically insignificant, indicating the absence of trends. [11]

[11] Another alternative estimator that could be suitable for my data set-up is the TWFE stacked DiD model (Cengiz et al., 2019) implemented in Vu (2019). All of these proposed models by Callaway and Sant'Anna (2020), Sun and Abraham (2020), and Cengiz et al. (2019) are sufficient to recover the true treatment effects in the presence of treatment effects heterogeneity. For the sake of brevity, I report estimators proposed by Callaway and Sant'Anna (2020) and Sun and Abraham (2021).

5.3 Heterogeneity in Program Effects

Tables 4 and 5 show the results from estimating equations (3)-(6) where the effect of the program on mortality outcomes varies by gender (column 1), by rural (column 2), by caste (column 3) and by SES (column 4). Table 4 reports the results on infant mortality and Table 5 reports the results on under-five mortality. In addition to the main and interaction effects, I also report the linear combination of main and interaction effects along with the associated standard errors and significance. Column (1) of Table 4 suggests that UIP did not have a different effects for boys and girls, i.e., infant mortality decreased by about the same amount for both girls and boys due to UIP (though the point estimate is negative for girls, suggesting that girl might have benefited a little more). Column (2) suggests that the program had a null effect in urban areas (the point estimate is 0.05 percentage points and is statistically insignificant) and rural areas had a negative effect that is significantly different both from the urban effect and from zero at a 5 percent level of significance. That a reduction in child mortality is found only in rural areas may be due to urban people already having access to vaccinations even before UIP or just that health conditions and care tends to be better in urban areas (as indicated by the much lower infant mortality rates at the outset). Thus, there is more room for improvement in the rural areas, and that is where I find mortality effects. In column (3), there is evidence of differential effects by caste–the program effects are larger for SC and OBC groups compared with general caste groups. The effect for ST children is negative but not significant.[9] Finally, the program also has negative and significant impact on infant mortality for children from poor households (column 4). There was no reduction in mortality among children from middle- and high-income households. Columns (3) and (4) both inform on whether the effects of UIP vary by household socioeconomic status (there is much overlap between being in a low social group and being poor), and they both clearly show that it is the children from disadvantaged households who experience the declines in mortality; there is no effect of UIP on higher caste groups or the middle-income and rich. Children from a higher caste and non-poor households have lower mortality rates at the outset, and may have been getting vaccinated to some extent already.

[9] Similar to the SC and the OBC, the ST are also a disadvantaged group with a higher-than-average child mortality so perhaps it is surprising that UIP did not reduce this group's child mortality more. This weak effect is probably due to the fact though UIP was far-reaching, it may not have reached many of the places where the ST reside; the ST mostly live in remote, sparsely settled rural areas.

I show the results for under-five mortality in Table 5. They are similar to the infant mortality results in Table 4 except that the estimated reduction in mortality is higher for reasons stated earlier (the under-five mortality measure includes mortality of infants and older children). Similar to infant mortality results, it appears that children from rural, lower caste groups, and poor households are benefiting from UIP in terms of increased survival rate.

5.4 Placement of the UIP program

Although the UIP placement is not completely random, I address this issue in my paper in the following three ways– (a) Institutional Details (b) Regression estimate, (c) control for fixed characteristics of the districts

The first strand of evidence to address the nonrandom placement of the program relies on the institutional details collated from the various government sources. India was a signatory of the WHO-UNICEF Alma-Ata declaration in 1978 in which "Health for All" was proposed as the fundamental human right. After this declaration WHO and UNICEF partnered together to launch Universal Immunization Program in WHO member countries. There was an international movement for launching the Universal immunization program in WHO member countries during the 1980s. India in partnership with UNICEF launched the Universal Immunization Program (UIP) in 1985-86. In the above context, I argue that the launch of UIP in India was exogenous to the prevailing disease environment in the country and it was an outcome the of international availability of funds and vaccines. Since it was a centrally sponsored program, the decision to implement the program across districts was taken either by the federal government or the state government and it is very unlikely that local factors influenced which district will get the program first. It is also true that there was resource and funds constraint and it was nearly impossible to implement the program in all the districts in one go and this led to phase-in of the program across districts.

Second, I investigate this issue in a regression framework. I use the "pre-program" characteristics of the districts to predict the rollout of the program. The goal is to check the existence of any systematic link between district-level observed characteristics and the rollout out year of the program. Results are reported in Table 6. Overall, none of the pre-UIP district-level variables seem to predict the year in which a district adopted UIP.

Third, I use the RCH survey data and construct district-wise average mortality for the "pre-program" period (1980-1984). I ranked the states and the districts based on the average child mortality rate. National Rank is the rank of districts based on average mortality across India and district rank is the rank within a particular state. Rank is in ascending order implying that higher average mortality districts have a higher rank. I run an ordered logit regression of the UIP implementation year on the "pre-program" average mortality of the district. The dependent variable is the year of the district's UIP start date expressed as an index equal to 1 in 1986, 2 in 1987, and so on. I check the systematic relation between program implementation year and district-level average mortality by estimating the following equation:

$$Y_{ds} = \beta_0 + \beta_1 NationalRank_{ds} + \gamma_s + e_{ds} \tag{7}$$

$$Y_{ds} = \beta_0 + \beta_1 DistrictRank_{ds} + \gamma_s + e_{ds} \tag{8}$$

Where Y_{ds} is the UIP year for an district d in state s. γ is state fixed effect, and e_{ds} is the independently-distributed individual error term. State fixed effects control for any unobserved time-invariant state characteristics. The coefficient of interest is β_1. Results in appendix Table 4 show that disease burden in the district as reflected in the district-level mortality rate did not influence the placement of UIP across districts (Table 4, appendix).

Fourth, for my identification, it is not crucial for the program placement to be random. To address the concern that program placement might be a function of some district-specific fixed characteristics, I include the district fixed effect in my base regression. I also include the year of birth fixed effect to control for time-variant cohort-specific characteristics.

Finally, resources made available to districts under UIP may be correlated with the institutional quality and capacity of the district–the quality of health infrastructures may draw more or fewer resources and may affect mortality outcomes. In this case, the UIP's effects could be overestimated if the districts that experienced a greater jump in immunization rate improved their overall basic health service provision during this period. We argue that the inclusion of district-fixed effects and the results in Tables 2 and 6 showing insignificant predictors of UIP placement should be able to address these concerns.

5.5 Testing for Differential Trends in Child Mortality

The identifying assumption for my empirical strategy is that in absence of the program, the difference in outcomes of younger and older cohorts would be the same between earlier-implementing districts and later-implementing districts. We would not be able to interpret the coefficients for "Exposed to UIP" in Table 3 as the causal effect of the immunization program if the aforementioned were not true. In this subsection, I assess the validity of this assumption by performing two "placebo tests": (1) estimating equation (2) using only older cohorts that have never been exposed to UIP but where I falsify their treatment status; and (2) estimating equation (2) using a series of outcome variables that are unlikely to be affected by the program.

To test the possibility of a differential trend in infant and child mortality, I do placebo tests in which I only use data on older cohorts who are not affected by the program (Duflo, 2004; Angrist, Chin, and Godoy, 2008). I take advantage of older cohorts born between 1977 and 1982 who are not treated.[10] In this placebo test, I assign pseudo-treatment to cohorts born during 1977-1982 as if the program were implemented during 1977-1982 (instead of 1985-1990). For example, districts that got the program in 1985-1986, now get the pseudo program in 1977-78. For these cohorts, I run the same basic specification outlined in equation (2). Columns (1)-(2) of Table 7 reports the results of this placebo test. The outcome variables are Pr(Infant Mortality) and Pr(Under-Five Mortality). The coefficient for the "Exposed to UIP" variable should be zero if the identifying assumption is correct, i.e., the cohort change in mortality would have been the same in earlier-implementing districts and later-implementing districts in the absence of UIP. This is because nobody in the sample actually got exposed to UIP. In Columns (1)-(2), I find statistically insignificant coefficients for the pseudo-treatment variable, and point estimates are actually close to zero and positive. This suggests that district-specific cohort trends do not appear to be confounding the estimates of the effect of UIP using my difference-in-differences approach and that the difference-in-differences estimates shown in Tables (3)-(5) can be interpreted as the causal effect of UIP.

[10] I am unable to use older cohorts born before 1977 because of data limitations. There are very few observations for the period before 1977 because of the sampling nature of the RCH survey. The survey has information on women who are 15-44 years old at the time of the survey (2002-04). Given the age restriction on the mothers, there are very few children born before 1977.

Another placebo test that can be performed is to check outcomes that are unrelated to UIP itself but are susceptible to the same unobserved district-cohort changes that are putatively confounding the difference-in-differences estimates. For example, one might be concerned that earlier-implementing districts had more improvements in health care which in turn is causing a greater decline in infant and child mortality compared to later-implementing districts. Of course, the previous placebo test using older cohorts does not support this claim, but arguably the differential trends estimated using those cohorts do not apply to the cohorts in my main analysis because of some regime shift. The RCH survey does not collect many health outcomes across people of different age groups, but it does have information on three health outcomes–blindness, tuberculosis (TB), and malaria. I perform placebo tests in which I examine the impact of the program on blindness, TB, and malaria. I dichotomized these three outcomes.[11] The medical literature suggests that immunization shots given under UIP are unlikely to protect children from blindness, TB, or malaria. If this is true, I should not observe any correlation between exposure to UIP and these three health outcomes. Columns 3-5 of Table 7 report the results. The coefficients for "Exposed to UIP" is not significantly different from zero for any of the health outcomes. Thus, it does not appear that earlier-implementing districts have significantly more progress in health care compared to later-implementing districts. If anything, it is the other way around since the point estimate is positive for blindness and malaria (and for this outcome as for mortality, a more positive number means worse health). The result of this placebo test provides more support for the validity of the identifying assumption for the difference-in-differences approach that I use to estimate the causal effect of UIP.

5.6 Discussion of Child Mortality Results

I find that on average, UIP reduced infant mortality by 0.4 percentage points and under-five child mortality by 0.5 percentage points. The effects are more pronounced in rural areas, for poor people, and for members of historically disadvantaged groups. For example, children born into poor households were 0.9 percentage points less likely to die within the first twelve

[11] The survey asks the blindness question for each member of the household in the household roster. The answers are categorized as partially blind, completely blind, night blind, and not blind. I construct an outcome variable "blindness" by combining partially blind, completely blind, and night blind together.

months and 1.3 percentage points less likely to die within the first five years. Thus, there are huge benefits of a mass immunization program in terms of reducing child mortality.

As mentioned before, these estimates may underestimate the true effect of UIP for a couple of reasons. First, because the consumption of vaccines has a positive externality, the control group may benefit indirectly from UIP. Second, an immunization program may provide more benefits in urban areas than the estimated effects in this study, because some children in urban areas might have been vaccinated under the EPI program (before UIP). Finally, I also report DiD estimates with state-specific time trends (Table 5 in the appendix). The estimates are qualitatively similar but are imprecisely estimated and thus reflect that the inclusion of state-specific time trends make the DiD estimates sensitive.

The significant effects on child mortality are consistent with children's immune systems being strengthened. This would suggest that there would be significant decreases in child morbidity, too, due to UIP. Immunization not only protects against specific diseases (whose symptoms may be deadly to infants but only illness among older children) but also improves the overall immune system of the body (which protects the children from other illnesses). It is likely that the benefits of UIP on child health conditional on surviving would not be confined to children from a rural, low caste, and poor households; children from urban areas, higher caste, and non-poor households may be far from the margin of survival but there is still room for improvement in terms of health status. Unfortunately, I do not have the data to directly assess the effects of UIP on child health outcomes besides the extreme outcome of mortality.[12]

It is worth highlighting that UIP has been a successful program in reducing child mortality in India even though India is characterized by poor service delivery mechanisms and high absenteeism of health staff. Banerjee, Deaton, and Duflo (2004) discuss a very bleak picture of public and private healthcare provision in the Udaipur district of India and find that 45% of medical personnel are absent in health subcenters. A similarly high level of absenteeism has also been found in a nationally representative survey of primary health centers in India by Chaudhury et al. (2006). The fourth round of the Indian National Family Health Survey shows that about 51% of women perceived the unavailability of doctors as a big problem

[12]The RCH survey has extensive health measures for children under age 5. Given that the RCH survey is collected in 2002-04, all these children would have been exposed to UIP, leaving no variation in treatment. Thus these rich child health measures are not usable to estimate the impact of UIP.

for not seeking medical care in public health facilities. Despite these impediments, the immunization program did achieve its intended objective of reducing mortality among Indian children.

6 Effect of UIP on the Educational Outcomes of Surviving Children

Some recent studies using plausibly exogenous variation in child health from interventions to reduce worm diseases and malaria and from school nutrition programs have found a causal relationship running from child health to education (these were discussed in subsection 2.2). In particular, improving child health improves educational outcomes. The ability to attend school more regularly and often and to concentrate on studies better when one is healthier is thought to be responsible at least in part. My results show that UIP reduced child mortality and speculate that it reduced child morbidity too. Given these beneficial effects of UIP on child health, it is natural to ask what are the consequences for the educational outcomes of the surviving children.

6.1 Estimation Results

Education results are based on children who survived beyond age five. Table 8 presents the results of estimating equation (2) using each of the educational outcomes in turn-Pr(Literate), Pr(Primary School Completion), Pr(Middle School Completion), Pr(Secondary School Completion), and Years of Schooling.[13] Column (1) suggests that the program has no effect on the probability of being literate. The educational outcomes are conditional on the child being literate.[14] Column (2) suggests there is no significant effect of UIP on years of schooling completed. This masks a nonlinear effect of UIP on schooling. UIP significantly decreased the

[13]I estimate these models using OLS, so in the case of a dichotomous outcome I am using the linear probability model. For the dichotomous outcomes, I have also used the logit model and found qualitatively similar results; these results are available upon request.

[14]The RCH survey asks for years of schooling completed only for people responding affirmatively to being literate. In theory, this could lead to selection bias when I examine the impact of UIP on years of schooling and the primary, middle, and secondary school indicator variables since the sample is conditional on being literate. Therefore, I use the Heckman two-step correction method to correct for the sample selection bias. Two steps include the selection equation and the outcome equation, In step 1, I estimate the selection model using the probit model–the dependent variable (being literate) as a function of the child (gender, age), mother's (age, education), and household characteristics (religion, caste). The selection equation is used to derive the inverse mills ratio (lambda). The inverse mills ratio derived from the selection equation is included as an explanatory variable while analyzing the UIP effect on educational outcomes. In practice, the bias on the estimated effect of UIP is negligible given the insignificant result in Column (1).

probability of primary school completion (by 4.7 percentage points), had no impact on middle school completion, and significantly increased the probability of secondary school (by 1.9 percentage points). Furthermore, educational outcomes likely follow a different cohort-specific trend. To allow for this cohort-specific trend, I report the DiD estimates with state-specific time trends are reported in Table 6 in the Appendix. The results are qualitatively similar to the findings in Table 8.

Results in Table 8 suggest that though UIP did not raise years of schooling on average, it reduced schooling at low levels of education and increased it at higher levels of education. To get more detail on the effect of UIP at different points in the education distribution, I estimate equation (2) for each level of schooling k, where the dependent variable is Pr(Years of Schooling $\geq$ k), where k = 1 to 15. The estimated coefficients with the 95-percent confidence interval are plotted in Figure 1. Each point on the graph is from a different regression. The “S" shape of Figure 1 suggests that the program decreased the number of years of primary schooling for some children, but increased the number of secondary schooling for others. At the low end, there is a shift away from completing 5 years of schooling toward completing 2-4 years of schooling. At the high end, there is a shift away from completing fewer than ten years of schooling toward completing 10-12 years of schooling. Apparently, the decrease at the lower end of the education distribution offsets the gains at the upper end, leading to a zero average effect on years of schooling.

In Figure 2, I perform the same analysis as in Figure 1 but separately for children in rural and urban areas. Comparing Panels A and B, we see that all the negative impact of UIP at the low end of the education distribution is coming from rural areas-the graph in Panel B is flat at zero for the first nine years of schooling. This is especially interesting since it was only in rural areas where UIP had an impact on child mortality. UIP increased schooling at the high end of the education distribution for children in both the urban and rural areas, however.

In Figure 3, I perform the same analysis as in Figure 1 but separately for children from low, middle, and high socio-economic status households. We see that all the negative impact of UIP at the low end of the education distribution is coming from the poor households–the graph in Panels B and C does not have the trough at the lower levels of schooling. This is also interesting because it was only in poor households where UIP significantly reduced child mortality. The effects have an inverse U-shape at the high end of the education distribution in all three wealth categories though results are only significant for the poor category.[15]

6.2 Channels for the Effect of UIP on Educational Outcomes

UIP had mixed impacts on children's educational outcomes. It appears to have decreased the number of primary grades completed for some children but increased the number of secondary grades completed for other children. This nonlinear effect in which some children have worse educational outcomes and others have better ones is unusual vis-a-vis the existing literature which has tended to find positive effects on education for health interventions that improve child health. There are various explanations for the results that I find in Table 8 and Figures 1-3. Though I cannot conclusively pin down the pathways, here I describe several hypotheses consistent with the results. These hypotheses are not mutually exclusive, and may each have a part in the overall results.

6.2.1 Composition Effect

Change in the composition of the surviving pool of children can be one factor that is driving the nonlinear effects on education. First, UIP reduced child mortality. I will call the children who UIP saved from dying marginal children. Though they are alive (and in this sense, have better health than those without UIP), these marginal children are probably less healthy than the average child. Because they have worse health than surviving children, they have worse educational outcomes. This may be because they attend school less frequently, have less capacity to focus and learn, or take longer time to complete normal tasks.

As argued earlier, UIP also likely reduced child morbidity among inframarginal children (i.e., the children far from the margin of survival). Since their health is better in the traditional sense (i.e., as in the other papers estimating the causal impact of child health on education such as Miguel and Kremer, 2004 and Bleakley, 2007), we might expect their educational outcomes to improve as in these other papers.

[15]The lack of significance is likely because the number of observations is much less when I perform the analysis separately for each SES category; results remain significant in Panel A because 45% of the sample is in the poor SES category.

That UIP improved health on two margins, from dying to survival for the children on the margin of survival, from less healthy to more healthy for inframarginal children provides a cohesive story for the nonlinear effects on education. On the one hand, the marginal children are likely to be concentrated on the lowest parts of the education distribution, causing there an estimated reduction in primary school completion. Corroborating this assertion is that in Figures 2 and 3, we observe a negative impact at the low end of the education distribution only among those children who experienced a decline in mortality-the rural and the poor. On the other hand, the inframarginal children are on higher parts of the distribution and their improvement in health causes them to attain more years of schooling.

6.2.2 Quantity-Quality Tradeoff

Quality-quantity (QQ) trade-off could be another explanation for mixed results on education (Becker, 1960; Becker and Lewis, 1973; Becker and Tomes, 1976). The QQ model implies that an increasing marginal cost of quality (child outcome) with respect to quantity (number of children) leads to a tradeoff between quantity and quality. In this paper, I find that UIP reduces child mortality. This would have increased the number of surviving children, though I do not test this in this study. In a companion paper, I examine the impact of UIP exposure on the fertility of women and the number of surviving children (Kumar, 2009). Kumar (2009) shows that women whose first-born child was exposed to the UIP program had a lower likelihood of subsequent and cumulative fertility and the birth interval between the first and second births also increased due to UIP. Empirical studies generally conclude that child mortality reduction modestly decreases the number of births, increases the number of surviving children, and stimulates population growth.[18] In such cases of an increase in family size, parents have fewer resources to spend on each child, leading to less investment in the quality (education) of the children. Thus, the quantity-quality tradeoff explains the negative effect on primary school completion.

The empirical evidence on the quantity-quality tradeoff is mixed (e.g., Black et al., 2005; Angrist et al., 2005). It is possible that this tradeoff does not exist in developed countries

[18] Azarnert (2006); see Preston (1978) for a collection of demographic essays that come to such conclusion and Palloni and Rafalimanana (1999) for a broad survey of literature; see also Rutstein (1974), Chowdhary et al. (1976), Balakrishnan (1978), Olsen (1980), and Olsen and Wolpin (1983).

where there exists a well-structured public education system and welfare programs targeting childbearing and child care. In a developing country like India, the cost of child quality is mostly borne by the parents because these countries lack a well-functioning public education system and they do not have any support for childbearing and childcare. Thus, the quantity-quality tradeoff is more likely to be relevant to developing countries like India. Kugler and Kumar (2017) show the existence of a quantity-quality trade-off in India, particularly among low-wealth households. Li, Zhang, and Zhu (2008) find similar results in China–negative correlation between family size and children's education.

6.2.3 Other Explanations

School quality is another channel that can drive the nuanced education results in the paper. The reduction in child mortality rates caused by UIP may have contributed to an increased number of surviving children, subsequently leading to crowded classrooms and exerting a downward influence on the standard of instruction. Classroom overcrowding and worsening student-teacher ratio in developing countries can have a detrimental impact on the quality of education. Prior studies show that a large enrollment increase in primary schools worsened the quality of education in developing countries (Pritchett 2013, World Bank 2018).

This is possible if the government is hard-pressed for resources and prioritizes other programs such as child health programs, potentially straining available resources for the education sector. However, a simultaneous failure to enhance the existing school infrastructure and educational quality can exacerbate the consequences of crowded classrooms. The resultant decline in school quality or prevalence of overcrowded classrooms may further manifest in diminished school participation or reduced learning outcomes. [12]

Another explanation for the negative effect on primary school completion is that improving child health may increase children's labor force participation. It is quite likely that when returns to schooling are low or if the family is credit-constrained, children join the agricultural field with their parents to support the family or they enter the child labor market. If a child is somewhat healthy but not so healthy that he can earn much working, he may be sent

[12] For example, the student-teacher ratio at the primary level is 33 students per teacher, while the world average is 21.75 students per teacher. However, the evidence on the impacts of class size reduction on learning outcomes is mixed—the typical findings are close to zero or very small (Angrist and Lavy, 1999; Datta and Kingdon, 2023).

to school. When his health improves say due to UIP, he may become healthy enough to work and therefore drop out of school.

Besides the story where the health of the inframarginal children improves, another story for why there may be positive effects on education is that UIP has increased the life expectancy of children, causing parents to invest more in children's education because there is a longer period to collect the returns. Jayachandran and Lleras-Muney (2009) find empirical evidence in support of this hypothesis. This hypothesis is unlikely to apply in the present case because UIP primarily affects the mortality of very young children. Conditional on surviving to age 5, children's life expectancy does not differ much with or without UIP. Yet, parents are likely not making educational investment decisions before age 5 in most of India and given this, UIP should not impact education through this mechanism. However, vaccinations indeed reduce morbidity at later ages even if mortality is largely unaffected, and this can add up to meaningful differences in productive days between UIP-exposed children and other children.

7 Conclusion and Policy Implications

By using the phase-in feature of India's Universal Immunization Program immunization program and eligibility rules that only granted vaccinations to children up to twelve months, I estimate the causal effect of UIP on children's health and education outcomes. I find that the program significantly reduces infant mortality and under-five mortality in India. Contrary to the popular belief that LMICs are plagued with inefficient program implementation capacity and poor public health service delivery system, this paper establishes that UIP successfully achieved its objective of reducing mortality in the context of a developing country.

Among surviving children, UIP had a negative impact on primary school completion for some, but a positive impact on secondary school completion for others. The results on education outcomes can be explained in terms of change in the composition of the surviving children due to the immunization program. The negative effect on education may be due to the lower quality of the "marginal child" similar to the argument made by Donohue and Levitt (2001) and Gruber, Levine, and Staiger (1999); UIP induced some children to survive who otherwise would have died, and these children may be less healthy. The negative results are also consistent with the quantity-quality trade-off where an unanticipated increase in household size due to the immunization program induces the households to under-invest in each child. On the other hand, the result that UIP increased the education of some children is likely due to improved health among those children who are not at the margin of survival.

The results of this paper have important policy implications for the design of optimal health and education policy in developing countries. While the program had the intended benefit of increasing the survival probability of young children, there are mixed results for educational outcomes with more children less likely to complete primary school. It may be that the resources of both families and schools were too severely constrained to meet the needs of the marginal children. A lesson may be that child health and education policies have to be considered jointly so that children not only survive but are also given adequate resources and opportunities to receive a decent education. Policymakers should provide additional resources to educate the marginal child and help them perform better. The provision of extra teachers, after-school instruction, supplemental instruction, or remedial education similar to "Balasakhi" would be a step forward in this direction (Banerjee et al., 2007; Glewwe and Muralidharan, 2016).

References

Alderman, Harold; Behrman Jere R.; Lavy, Victor and Menon, Rekha. "Child Health and School Enrolment: A Longitudinal Analysis." *Journal of Human Resources*, 2001, 36(1), pp. 185-205.

Almond, Douglas and Currie, Janet. "Killing me softly: the fetal origins hypothesis." *J. Econ. Perspect*. 2011, 25, 153–172.

Almond, Douglas; Currie, Janet and Duque, Valentina. "Childhood Circumstances and Adult Outcomes: Act II." *Journal of Economic Literature*, 2018, 56 (4): 1360-1446.

Angrist, Joshua D., and Victor Lavy. "Using Maimonides' rule to estimate the effect of class size on scholastic achievement." The Quarterly journal of economics 114, no. 2 (1999): 533-575

Angrist, Joshua; Lavy, Victor and Schlosser, Analia. "Multiple Experiments for the Causal Link between the Quantity and Quality of Children." 2010, *Journal of Labor Economics*, 28(4).

Angrist, Joshua; Chin, Aimee and Godoy, Ricardo. "Is Spanish-only Schooling responsible for the Puerto Rican Language Gap?" *Journal of Development Economics*, 2008, 85(1), pp. 105-128.

Annual Report. Ministry of Health and Family Welfare (MoHFW), 1987-88, Government of India.

Azarnert, Leonid V. "Child Mortality, Fertility, and Human Capital Accumulation." J*ournal of Population Economics*, 2006, 19, pp. 285-297.

Balakrishnan, T. R. "Effects of Child Mortality on Subsequent Fertility of Women in Some Rural and Semi-Urban Areas of Central Latin American Countries." *Population Studies*, 1978, 32, pp. 135-145.

Banerjee, Abhijit; Deaton, Angus and Duflo, Esther. "Wealth, Health, and Health Services in Rural Rajasthan." *American Economic Review*, 2004, 94(2), pp. 326-330

Banerjee, Abhijit; Cole, Shawn; Duflo, Esther and Linden, Leigh L. "Remedying Education: Evidence from Two Randomized Experiments in India." *Quarterly Journal of Economics*, 2007, 122(3), pp. 1235-1264.

Becker, Gary S. "An Economic Analysis of Fertility." Demographic and Economic Change in Developed Countries, Gary S. Becker, ed., Princeton: Princeton University Press, 1960.

Becker, Gary S. and Lewis, Gregg H. "On the Interaction between the Quantity and Quality of Children." *Journal of Political Economy*, 1973, 81(2), pp. S279-S288.

Becker, Gary S. and Tomes, Nigel. "Child Endowments and the Quantity and Quality of Children." *Journal of Political Economy*, 1976, 84(4), S143-S162.

Bertrand, Marianne; Duflo, Esther and Mullainathan Sendhil. "How Much Should We Trust Differences-in-Differences Estimates?" *Quarterly Journal of Economics*, 2004, 119(1), pp. 249-275.

Black, Sandra E.; Devereux Paul J. and Salvanes Kjell G. "The More the Merrier? The Effect of Family Size and Birth Order on Children's Education." *Quarterly Journal of Economics*, 2005, 120(2), pp. 669-700.

Bleakley, Hoyt. “Disease and Development: Evidence from Hookworm Eradication in the American South." *Quarterly Journal of Economics*, 2007, 122(1). pp. 73-117.

Bleakley, Hoyt. “Malaria Eradication in the Americas: A Retrospective Analysis of Childhood Exposure." *American Economic Journal-Applied Economics*, 2010, 2(2).

Bloom, David E.; Canning, David and Weston, Mark “The Value of Vaccination." *World Economics*, 2005, 6(3), July-Sep.

Bloom DE. “The value of vaccination." *Adv Exp Med Biol*. 2011; 697:1-8.

Bloom, David E., Canning, David & Shenoy, Erica S. “The effect of vaccination on children’s physical and cognitive development in the Philippines," *Applied Economics*, 2012, 44:21, 2777-2783, DOI: 10.1080/00036846.2011.566203.

Bobonis, Gustavo J.; Miguel, Edward and Sharma, Charu P. “Iron Deficiency, Anemia and School Participation." *Journal of Human Resources*, 2006, 41(4), pp.692-721.

Callaway, B Sant’Anna. P.H.C. Difference-in-Differences with multiple time periods." *Journal of Econometrics* Volume 225, Issue 2, 2021, Pages 200-230,

Chaudhury, Nazmul; Hammer, J.; Kremer, M. and Muralidharan, K.; and Rogers, H. “Missing in Action: Teacher and Health Worker Absence in Developing Countries." *Journal of Economic Perspectives*, Winter 2006, Vol. 20(1), pp. 91-116.

Chowdhary, A.K.M.A.; Khan, A. R. and Chen, L.C. “The Effect of Child Mortality on Subsequent Fertility in Pakistan and Bangladesh." *Population Studies*, 1976, 2, pp. 249-261.

Cengiz, D., A. Dube, A. Lindner, and B. Zipperer (2019). “The Effect of Minimum Wages on Low-Wage Jobs." The Quarterly Journal of Economics 134 (3), 1405–1454.

Clarke, P. “Introduction. In Education Reform and the Learning Crisis in Developing Countries (Cambridge Education Research, pp. 1-13)." Cambridge: Cambridge University Press. doi:10.1017/9781108973700.001

Currie, J., Vogl, T., 2013. “Early-life health and adult circumstance in developing countries". *Ann. Rev. Econ*. 5, 1–36.

Cutler, David; Fung, Winnie; Kremer, Michael, Singhal, Monica & Vogl, Tom “Early-Life Malaria Exposure and Adult Outcomes: Evidence from Malaria Eradication in India." *American Economic Journal: Applied Economics*, 2010, 2(2): 72-94.

Sandip Datta, Geeta Gandhi Kingdon. Class Size and Learning: Has India Spent Too Much on Reducing Class Size?, The World Bank Economic Review, Volume 37, Issue 1, February 2023, Pages 24–48, https://doi.org/10.1093/wber/lhac025.

Donohue, John J. and Levitt, Steve. “The Impact of Legalized Abortion on Crime," *Quarterly Journal of Economics*, 2001, 116(2), pp. 379-420.

Dow, W.H., Philipson, T.J., Sala-i-Martin, X., 1999. Longevity complementarities under competing risks. American Economic Review 89 (5), 1358–1371.

Driessen, J., Razzaque, A., Walker, D. & Canning, D. “The effect of childhood measles vaccination on school enrolment in Matlab, Bangladesh", *Applied Economics*, 2015 47:55, 6019-6040, DOI: 10.1080/00036846.2015.1061647.

Glewwe, Paul and Kremer, Michael. "Schools, Teachers, and Education Outcomes in Developing Countries." Handbook of the Economics of Education, Volume 2, edited by Eric A. Hanushek and Finis Welch. New York, NY: Elsevier, 2006, pp. 945-1017.

Glewwe, Paul and Muralidharan, Karthik "Chapter 10 - Improving Education Outcomes in Developing Countries: Evidence, Knowledge Gaps, and Policy Implications," Editor(s): Eric A. Hanushek, Stephen Machin, Ludger Woessmann, Handbook of the Economics of Education, Elsevier, Volume 5, 2016, Pages 653-743.

Gruber, Jonathan; Levine, Phillip B. and Staiger, Douglas. "Abortion Legalization and Child Living Circumstances: Who is the Marginal Child?" *Quarterly Journal of Economics*, 1999, 114(1), 263-291.

Gupta, J.P. and Murali, Indira. "National Review of Immunization Programme in India." 1989, MoHFW, National Institute of Health and Family Welfare, New Delhi.

Hamory, J; Miguel, E.; Walke,r M.; Kremer, M. & Baird, Sarah "Twenty-Year economic impacts of deworming." *Proc Natl Acad Sci* USA 2021;118. doi:doi:10.1073/pnas.2023185118.

International Institute for Population Sciences. "National Family Health Survey (NFHS-4) 2015–2016: India. 2017Mumbai: IIPS.

Jayachandran, Seema and Lleras-Muney, Adriana. "Life Expectancy and Human Capital Investments: Evidence from Maternal Mortality Declines." *Quarterly Journal of Economics*, 2009, 124(1), 349-397.

Kane, M. and Lasher, H. "The Case for Childhood Immunization." Occasional Paper No. 5, Children's Vaccine Program at Path, Seattle, WA, 2002.

Koeing MA; Khan MA; Wojtynink B.; Clemens J.D. "The impact of measles vaccination upon childhood mortality in Matlab, Bangladesh." *Bulletin of the World Health Organization*, 1990, 68, pp. 441-447.

Krämer, Marion, Kumar, Santosh, and Vollmer, Sebastian. "Improving Child Health and Cognition: Evidence from a School-Based Nutrition Intervention in India," *The Review of Economics and Statistics*, (2021), 103(5): 818–834.

Kremer, M., Brannen, C., Glennerster, R. "The challenge of education and learning in the developing world," *Science*, 340 (2013), pp. 297-300, 10.1126/science.1235350.

Kugler, Adriana and Kumar, Santosh "Preference for Boys, Family Size, and Educational Attainment in India," *Demography*, 2017, 54, 835-859.

Kumar, Santosh. "Fertility and birth spacing consequences of childhood immunization program: Evidence from India." MPRA Paper No. 31807.

Li, Hongbin; Zhang, Junsen and Zhu, Yi. "The Quantity-Quality Tradeoff of Children in a Developing Country: Identification Using Chinese Twins." *Demography*, 2008, 45(1); 223-243.

Lucas, Adrienne M. "Malaria Eradication and Educational Attainment: Evidence from Paraguay and Sri Lanka." *American Economic Journal: Applied Economics*, 2010,2 (2): 46-71.

Miguel, Edward, and Kremer, Michael. "Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities." *Econometrica*, 2004, 72(1), pp.159-217.

Nandi, A., Kumar, S., Shet, A., Bloom, David E., and Laxminarayan, R. "Childhood vaccinations and adult schooling attainment: Long-term evidence from India's Universal Immunization Programme," *Social Science & Medicine*, Volume 250, 2020, 112885, ISSN 0277-9536, https://doi.org/10.1016/j.socscimed.2020.112885.

Olsen, R. J. "Estimating the Effect of Child Mortality on the Number of Births." *Demography*, 1980, 17, pp. 429-443.

Olsen, R. J. and Wolpin, K.I. "The Impact of Exogenous Child Mortality on Fertility: A Waiting Time Regression with Dyanamic Regressors". *Econometrica*, 1983, 51, pp. 731-749.

Oster, Emily. "Does Increased Access Increase Equality? Gender and Child Health Investments in India." *Journal of Development Economics*, 2009, 89(1):62-76.

Ozier, Owen. "Exploiting Externalities to Estimate the Long-Term Effects of Early Childhood Deworming." *American Economic Journal: Applied Economics*, 2018, 10 (3): 235-62.

Palloni, A. and Rafalimanana, H. "The Effect of Infant Mortality on Fertility Revised: New Evidence from Latin America." *Demography*, 1999, 36, pp. 41-58

Preston, S. H. "The Effect of Infant and Child Mortality on Fertility." Academic, New York, 1978.

Pritchett, Lant. "The Rebirth of Education: Schooling Ain't Learning." Brookings Institution Press, 2013. JSTOR, http://www.jstor.org/stable/10.7864/j.ctt1gpccb8. Accessed 27 Mar. 2024.

Rutstein, S. O. "The Influence of Child Mortality on Fertility in Taiwan." *Studies in Family Planning*, 1974, 5, pp. 182-188

Sathyamala, C. "Immunization, The Technology Missions." *Seminar*, 1989, 354, New Delhi, February.

Sun, Liyang, and Sarah Abraham. 2021. "Estimating Dynamic Treatment Effects in Event Studies with Heterogeneous Treatment Effects." *Journal of Econometrics* 225 (2): 175–99.

UNESCO Institute for Statistics (UIS) "More than One-Half of Children and Adolescents are Not Learning Worldwide." 2017, UIS fact sheet No. 46. Montreal: UIS. http://uis.unesco.org/sites/ more-than-half-children-not-learning-en2017.pdf.

UNICEF. "Coverage Evaluation Survey." 2002, Government of India, Ministry of Health & Family Welfare and UNICEF.

United Nations Inter-agency Group for Child Mortality Estimation (UN IGME). "Levels and trends in child mortality, United Nations Children's Fund," 2019, https://data.unicef.org/resources and-trends-in-child-mortality/.

Vu, Khoa. "Higher Education Expansion and the Rise of the Service Economy in Vietnam". Working paper, 2019.

WHO (http://www.who.int). World Health Organisation, 2021. Immunization coverage [WWW Document]. URL https://www.who.int/news-room/fact-sheets/detail/immunization-coverage (accessed 10.23.21).

World Bank. "World Development Report: Realizing the Promise of Education for Development." 2018, Washington, DC, World Bank

Table 1: Descriptive Statistics

	Sample used in Child Mortality Analysis		Sample used in Education Analysis	
Variables	Mean	S.D.	Mean	S.D.
Treated	0.75	(0.43)	0.67	(0.47)
Rural	0.69	(0.46)	0.68	(0.46)
Female	0.48	(0.50)	0.50	(0.50)
Low SES	0.47	(0.50)	0.45	(0.50)
Middle SES	0.31	(0.46)	0.32	(0.47)
High SES	0.22	(0.41)	0.22	(0.42)
ST	0.16	(0.37)	0.17	(0.37)
SC	0.17	(0.38)	0.17	(0.37)
OBC	0.38	(0.48)	0.37	(0.48)
Hindu	0.76	(0.43)	0.74	(0.44)
Muslim	0.13	(0.33)	0.14	(0.35)
Christian	0.07	(0.25)	0.07	(0.25)
Birth Order	2.37	(1.38)		
Mother Age	37.06	(3.95)		
Mother Literate	0.39	(0.49)		
Infant Mortality	0.09	(0.28)		
Under-five Mortality	0.11	(0.31)		
Literate			0.82	(0.39)
Years of Schooling			6.15	(3.39)
Primary School Completion			0.65	(0.48)
Middle School Completion			0.36	(0.48)
Secondary School Completion			0.18	(0.38)
Age			13.71	(4.17)
Number of States and UTs	35		35	
Number of Districts	561		561	
Number of Observations	297,385		898,789	

Notes: ST, SC and OBC are Scheduled Tribe, Scheduled Caste, and Other Backward Caste respectively. ST and SC are historically disadvantaged groups. SES is the socio-economic status of the households. Different samples are used for the child mortality analysis and education analysis because the data are from different file of the RCH survey. The paper uses information from the household file for education analysis and information from fertility file for child mortality analysis. Household file has information on all the individuals who live in the house, and from the household file one woman is selected (15-44 years) to be asked about her complete fertility history.

Table 2: Balance test at the district level

Variables	Early-UIP districts (1)	Late-UIP districts (2)	Difference (1)-(2) (3)	p-value (4)
District Literacy Rate	34.20	33.09	1.11	0.52
Average Size of Landholding (Ha)	2.37	2.60	0.23	0.35
Gross Irrigated Area as % of GCA	30.35	29.68	0.66	0.84
District Population (Log)	14.10	13.91	0.19	0.02
Number of Villages in the District (Log)	7.01	7.04	0.04	0.74
# of Villages Connected to Road	548.50	477.93	70.57	0.12
# Health Centers	31.43	27.61	3.82	0.21
# Dispensary	65.31	55.75	9.56	0.13
# Hospital	21.64	20.51	1.13	0.67
# Community Health Workers	63.93	60.98	2.95	0.88
Electrified village (%)	58.96	51.69	7.26	0.11
# primary schools	1026.67	1018.31	8.35	0.91
# middle schools	243.83	232.24	11.57	0.59
# high schools	92.30	83.28	9.01	0.33
Hindu population (%)	80	84	-4	0.19
Muslim population (%)	12	10	2	0.44
SC population (%)	18	17	1	0.55

Notes: Early-UIP districts received UIP in 1985 and 1986, while late-UIP districts received UIP after 1986. SC denotes scheduled caste social groups.

Table 3: Effect of UIP on Child Mortality

Independent variables	Infant Mortality			Under-five Mortality		
	(1)	(2)	(3)	(4)	(5)	(6)
Exposed to UIP	-0.009***	-0.022***	-0.004*	-0.012***	-0.028***	-0.005**
	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.002)
Controls		✓	✓		✓	✓
District Fixed Effects			✓			✓
Year of Birth Fixed Effects			✓			✓
N	297,385	297,385	297,385	296,511	296,511	296,511
R Square	0.0002	0.03	0.04	0.04	0.05	0.005

Notes: Each column is from estimating a separate linear probability model. Robust standard errors clustered at the district level are in parentheses. *Controls*: gender and birth order of the child, mother's education, mother's age at birth, religion, caste, wealth index, and rural. Survey year dummy was used. District sample weights were applied. * shows significance at 10 percent level, ** at 5 percent level and *** at 1 percent level.

Table 4: Heterogeneous Program Effects on Infant Mortality

Independent Variables	Infant Mortality (1)	(2)	(3)	(4)
UIP	-0.003	0.0005	0.003	0.003
	(0.002)	(0.003)	(0.003)	(0.003)
UIP * Female	-0.0004			
	(0.003)			
UIP + UIP * Female	-0.004			
	[0.003]			
UIP * Rural		-0.006**		
		(0.003)		
UIP + UIP * Rural		-0.006**		
		[0.003]		
UIP * ST			-0.005	
			(0.004)	
UIP * SC			-0.015***	
			(0.005)	
UIP * OBC			-0.009***	
			(0.003)	
UIP + UIP * ST			-0.005	
			[0.004]	
UIP + UIP * SC			-0.015***	
			[0.005]	
UIP + UIP * OBC			-0.009***	
			[0.004]	
UIP * Poor				-0.012***
				(0.003)
UIP * Middle				-0.001
				(0.003)
UIP + UIP * Poor				-0.012***
				[0.003]
UIP + UIP * Middle				-0.002
				[0.003]
Controls	✓	✓	✓	✓
District Fixed Effects	✓	✓	✓	✓
Year of Birth Fixed Effects	✓	✓	✓	✓
N	297,385	297,385	297,385	297,385
R Square	0.04	0.04	0.04	0.04

Notes: Each column reports coefficients from separate linear probability models. Robust standard errors clustered at district level are in parentheses. Poor and middle are a dummy indicating Controls: gender and birth order of the child, mother's education, mother's age at birth religion, caste, wealth index, and rural. Poor and middle are dummies indicating wealth groups. Scheduled Caste (SC), Scheduled Tribe (ST), and Other Backward Caste (OBC) are caste groups. * shows significance at 10-percent level, ** at 5-percent level and *** at 1-percent level.

Table 5: Heterogeneous Program Effects on Under-five Mortality

Independent Variables	Under-five Mortality (1)	(2)	(3)	(4)
UIP	-0.003	0.003	0.005	0.004
	(0.003)	(0.003)	(0.003)	(0.003)
UIP * Female	-0.002			
	(0.003)			
UIP + UIP * Female	-0.002			
	[0.003]			
UIP * Rural		-0.011***		
		(0.003)		
UIP + UIP * Rural		-0.011***		
		[0.003]		
UIP * ST			-0.008*	
			(0.005)	
UIP * SC			-0.016***	
			(0.005)	
UIP * OBC			-0.013***	
			(0.004)	
UIP + UIP * ST			-0.008	
			[0.004]	
UIP + UIP * SC			-0.016***	
			[0.005]	
UIP + UIP * OBC			-0.013***	
			[0.004]	
UIP * Poor				-0.017***
				(0.004)
UIP * Middle				-0.002
				(0.003)
UIP + UIP * Poor				-0.017***
				[0.004]
UIP + UIP * Middle				-0.003
				[0.003]
Controls	✓	✓	✓	✓
District Fixed Effects	✓	✓	✓	✓
Year of Birth Fixed Effects	✓	✓	✓	✓
N	297,385	297,385	297,385	297,385
R Square	0.05	0.05	0.05	0.05

Notes: Each column reports coefficients from separate linear probability models. Robust standard errors clustered at district level are in parentheses. Poor and middle are a dummy indicating Controls: gender and birth order of the child, mother's education, mother's age at birth religion, caste, wealth index, and rural. Poor and middle are dummies indicating wealth groups. Scheduled Caste (SC), Scheduled Tribe (ST), and Other Backward Caste (OBC) are caste groups. * shows significance at 10-percent level, ** at 5-percent level and *** at 1-percent level.

Table 6: District-level pre-program predictors of UIP, 1981 census

Explanatory Variables	Year of UIP launch (1)	(2)
District Literacy Rate	0.027	0.018
	(0.023)	(0.023)
Average Size of Landholding (Ha)	-0.067	0.0077
	(0.57)	(0.074)
Gross Irrigated Area as % of GCA	0.001	0.001
	(0.008)	(0.008)
District Population (Log)	-0.490	-0.508
	(0.307)	(0.303)
Percent of Villages Connected to Road	-0.0001	0.000
	(0.0005)	(0.0006)
Number of Villages in the District (Log)	-0.323	-0.422
	(0.240)	(0.271)
# Health Centers		-0.015
		(0.012)
# Dispensary		0.004
		(0.004)
# Hospital		0.015
		(0.009)
# Community Health Workers		0.000
		(0.001)
State Fixed Effects	✓	✓
R Square	0.06	0.06

Notes: This table examines whether the program placement is based on pre-UIP district characteristics. District variables used in the estimation are from 1981 census (Pre-program period). Ordered logit coefficients are reported. Districts that received the program in 1985-86 are coded as 1, 86-87 as 2 and so on. GCA is Gross Cropped Area.

Table 7: Placebo tests

	Using Older Cohorts		Using other unrelated health measures		
	Infant mortality	Under-five mortality	Blindness	Tuberculosis	Malaria
Independent variables	(1)	(2)	(3)	(4)	(5)
Exposed to UIP	0.0003	0.0014	0.0009	-0.0002	0.00009
	(0.006)	(0.006)	(0.0007)	0.0003	(0.0006)
Controls	✓	✓	✓	✓	✓
District Fixed Effects	✓	✓	✓	✓	✓
Year of Birth Fixed Effects	✓	✓	✓	✓	✓
N	88,879	88,879	432,740	432,740	432,739
R Square	0.05	0.07	0.006	0.002	0.002

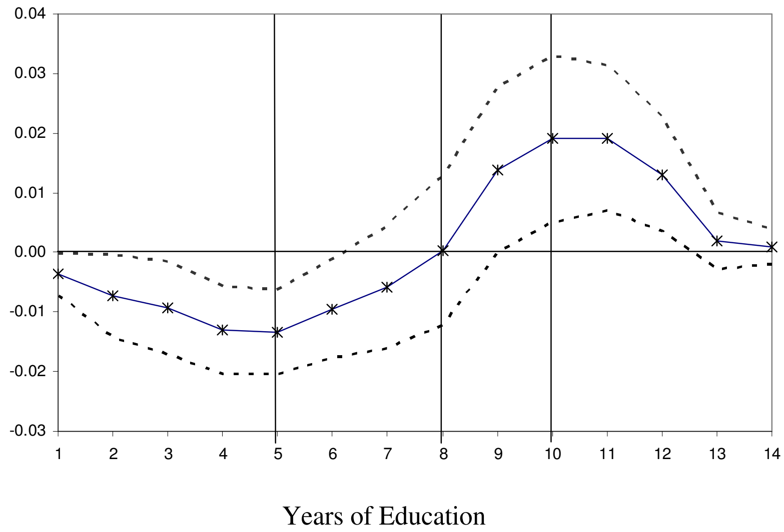
Notes: Each column shows LPM coefficients. Robust standard errors clustered at the district level are in parentheses. *Controls*: gender and birth order of the child, mother's education, mother's age at birth, religion, caste, wealth index, rural, and survey year dummies. District sample weights were applied.

Table 8: Effects of UIP on Education Outcomes

Independent Variables	Literate (1)	Years of Schooling (2)	Primary School Completion (3)	Middle School Completion (4)	Secondary School Completion (5)
Exposed to UIP	0.0011	0.006	-0.047***	0.0002	0.019***
	(0.003)	(0.028)	(0.004)	(0.006)	(0.007)
Controls	✓	✓	✓	✓	✓
District Fixed Effects	✓	✓	✓	✓	✓
Year of Birth Fixed Effects	✓	✓	✓	✓	✓
N	1099940	898789	898789	898789	898789
R-squared	0.18	0.65	0.55	0.47	0.33

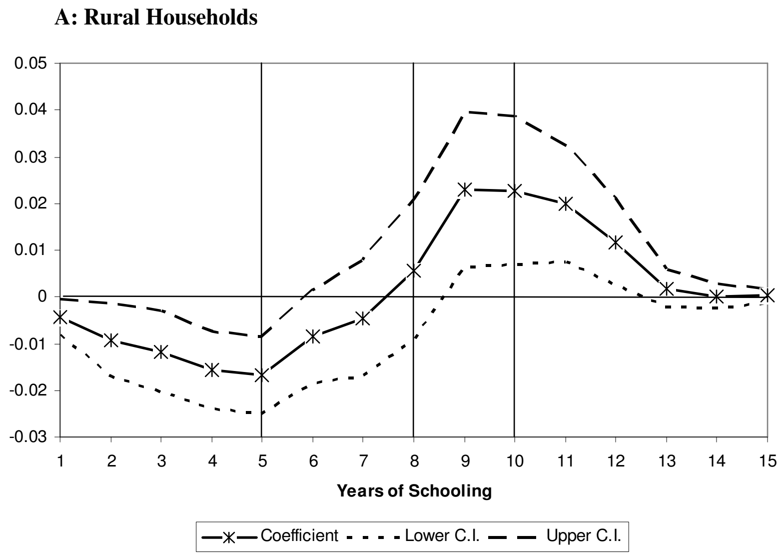
Notes: Columns (1), (3), (4), and (5) are from estimating a linear probability model. Columns (2), (3), (4), and (5) are from using Heckman. Two-Step Method to correct for selectivity-bias. Robust standard errors clustered at the district level are in parentheses. Poor is a dummy indicating household with low socio-economic status. Survey year dummy used. Scheduled Caste(SC) and Scheduled Tribe(ST)are traditionally disadvantaged minority group. OBC is Other Backward Caste. RCH district sample weights were applied. Primary School Completion is years of schooling $\geq$5, Middle School Completion is $\geq$8 and Secondary School Completion is $\geq$10. * shows significance at 10-percent level, ** at 5-percent level and *** at 1-percent level.

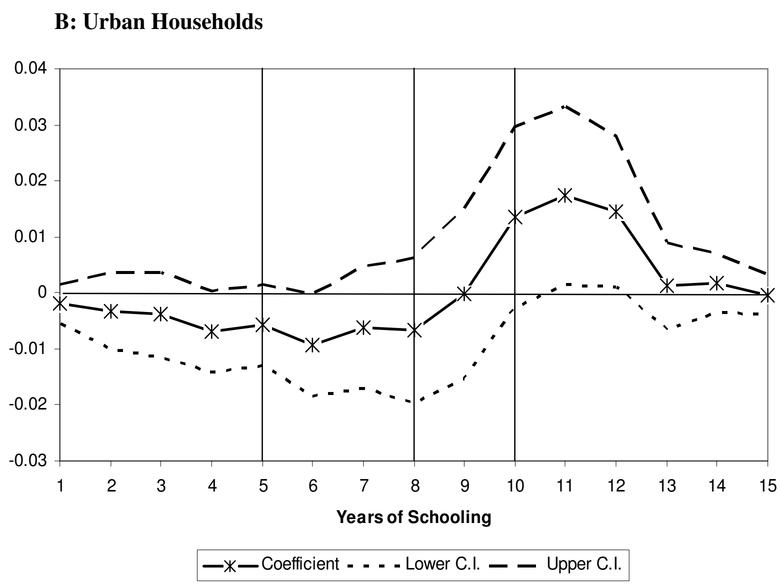
Figure 1: Effect of UIP at Each Years of Schooling



Notes: The figure shows difference-in-differences Pr(Years of Schooling $\geq$k) where k is labeled level of schooling. The broken line shows the 95-percent confidence interval. Each point is the estimated coefficient from a separate regression at each level of education. Primary school is completing grade five, middle school is completing grade eight and secondary school is completing grade 10.

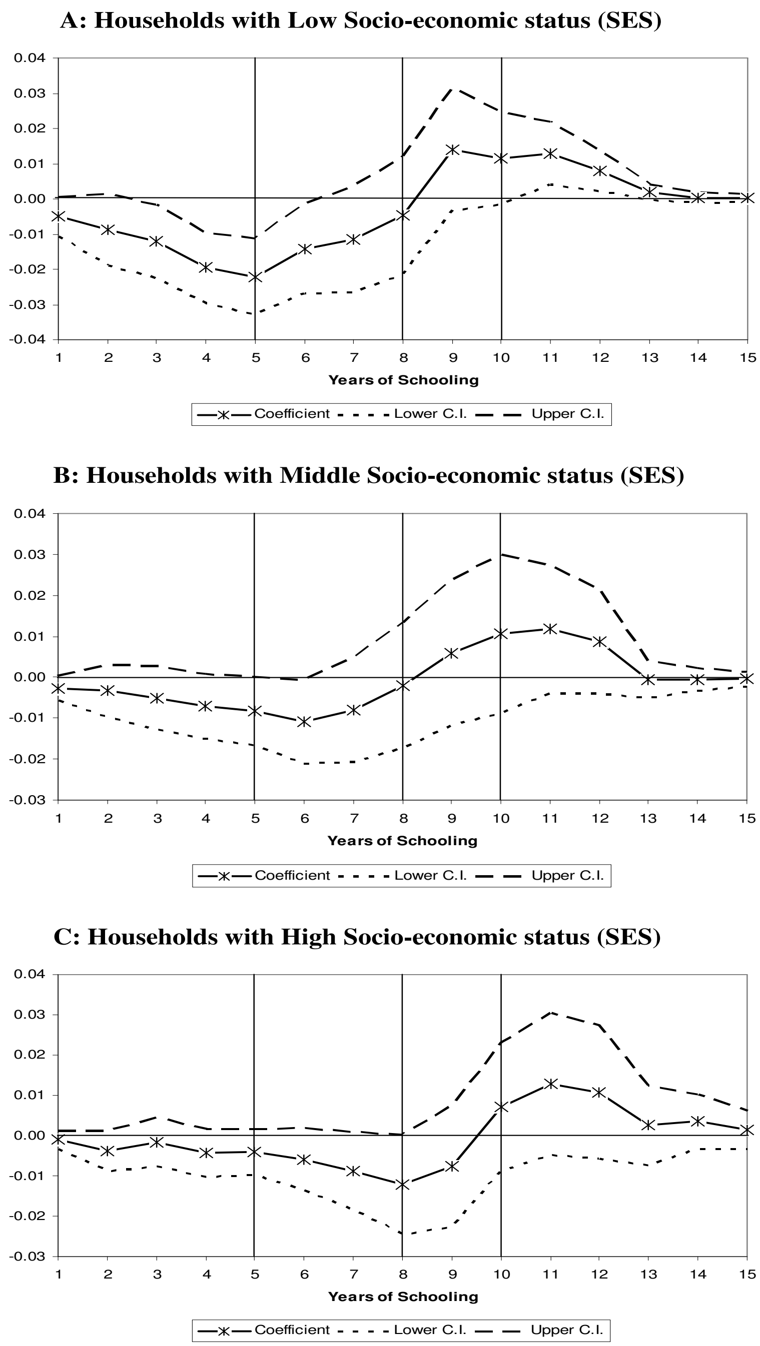
Figure 2: Heterogeneous Effects of UIP at Each Year of Schooling





Notes: The figure shows difference-in-differences Pr(Years of Schooling $\geq$k) where k is labeled level of schooling. The dashed line shows the 95-percent confidence interval. Each point is the estimated coefficient from a separate regression at each level of education.

Figure 3: Heterogeneous Effects of UIP at Each Year of Schooling



Notes: The figure shows difference-in-differences Pr(Years of Schooling $\geq$k) where k is labeled level of schooling. The dashed line shows the 95-percent confidence interval. Each point is the estimated coefficient from a separate regression at each level of education.

Appendix

Symptoms of the Disease

Table 1: Symptoms of the Disease

Diseases	Symptoms	Vaccines
Diphtheria	Sore throat and fever, Contagious Children under age 5 are at risk	DPT
Pertussis	Malnutrition, Pneumonia	DPT
Tetanus	Lock Jaw, Muscles Pain	DPT
Polio	Viral fever, Paralysis	Oral Polio
Measles	Skin rash, Running nose and Red eye	Measle
Tuberculosis	Chest pain, Coughing up blood Fever, Weight Loss and Appetite Loss Treatment takes 6-12 months	BCG

Immunization Schedule

Table 2: Immunization Schedule

	Age of the Child				
	At Birth	6 Weeks	12 weeks	16 Weeks	9-12 months
Primary Vaccination:					
BCG	Yes				
Oral Polio	Yes	Yes	Yes		
DPT		Yes	Yes	Yes	
Measles					Yes

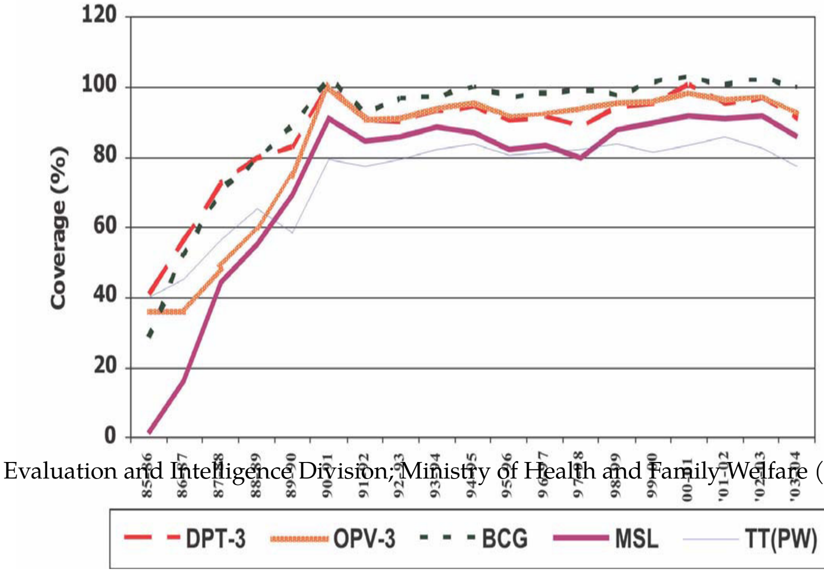
Exposure to the Program

Table 3: Exposure to the Program by Year of Birth and UIP Year

	UIP Year					
Year of Birth	1986	1987	1988	1989	1990	
1982						
1983						
1984						
1985	Yes					
1986	Yes	Yes				
1987	Yes	Yes	Yes			
1988	Yes	Yes	Yes	Yes		
1989	Yes	Yes	Yes	Yes	Yes	
1990	Yes	Yes	Yes	Yes	Yes	
1991	Yes	Yes	Yes	Yes	Yes	
Number of Districts	48	92	112	136	175	**563**

Table 3 explains the variations in exposure to the program across cohorts and districts. Staggered implementation of the program across districts between 1986-1990 led to differences in exposure to the program for children within the same birth cohorts. The official eligibility rule of getting all eight vaccines before the age of 12 months results in differences in exposure to the program across cohorts. Children are exposed to the program if their year of birth is either before or one year after their district of birth got the program. For example, children born in 1987 are treated if they were born in districts that got the program either before 1987 or one year after i.e. 1988 whereas children born in 1987 are considered untreated if they were born in districts that got the program in years later than 1988 (e.g. 1989, 1990). This illustrates that the same birth cohorts are exposed differently to the program depending on when their district got the program. This is a variation in exposure to the program across districts. Also, within a district, different birth cohorts are exposed differently to the program. Cohorts who are older than one year when the program was placed in their district are not exposed to the program whereas cohorts who are younger than one year when the program was placed in their district are exposed to the program. This is a variation in exposure to the program across cohorts.

Figure 1: Reported national vaccine coverage, by antigen from 1985 to 2004



Source: Evaluation and Intelligence Division, Ministry of Health and Family Welfare (MoHFW)

Figure 2: Phase-in of Districts over Years

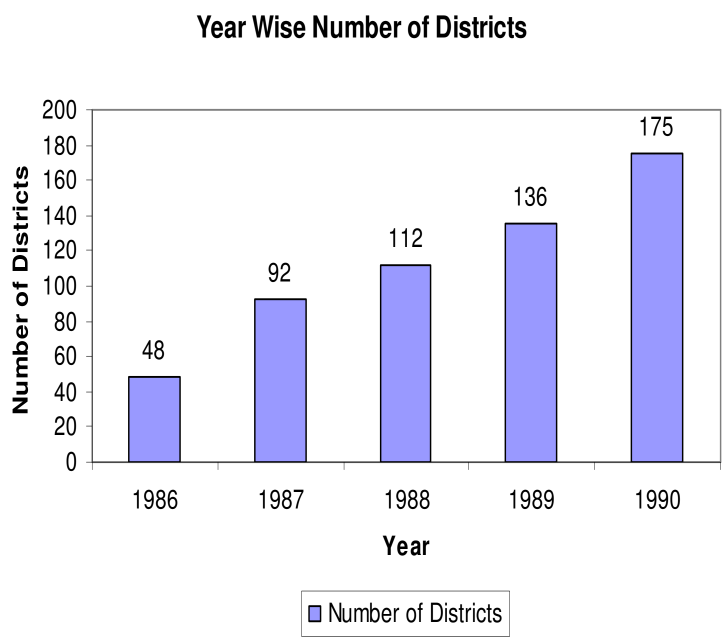


Figure 3: Effects of UIP on infant Mortality (Event study estimates)

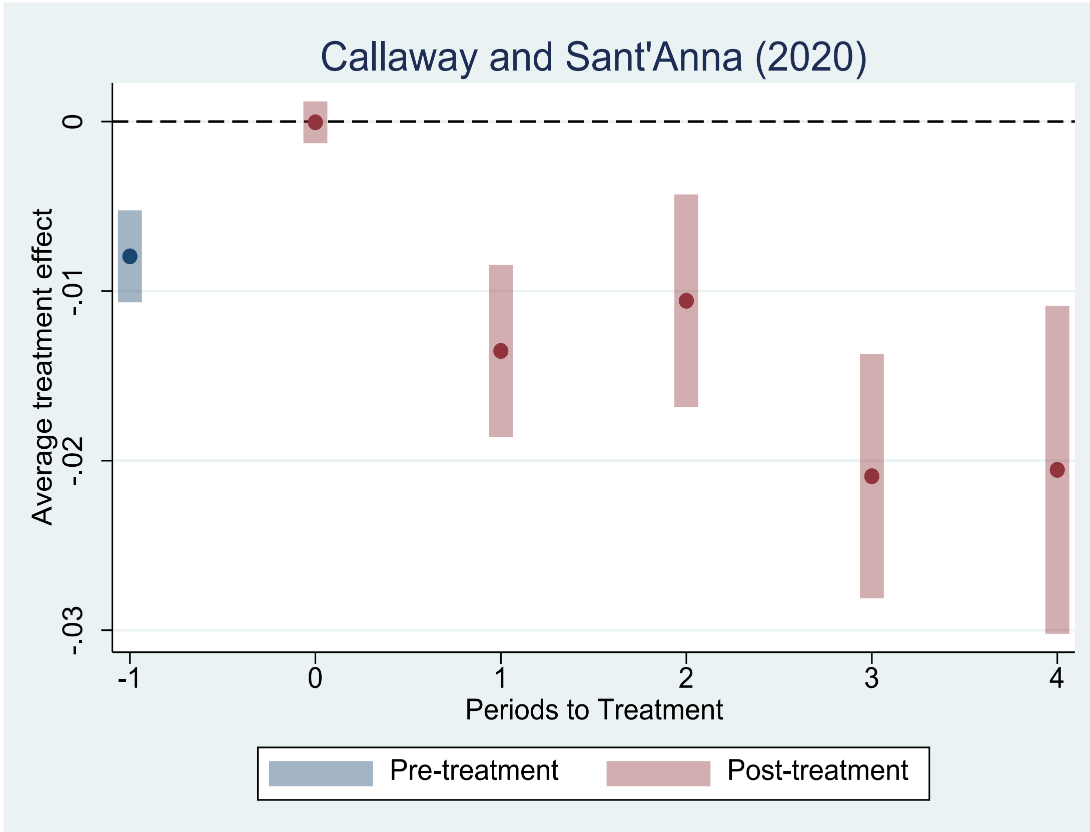


Figure 4: Pretrend (TWFE and Sun and Abraham (2021))

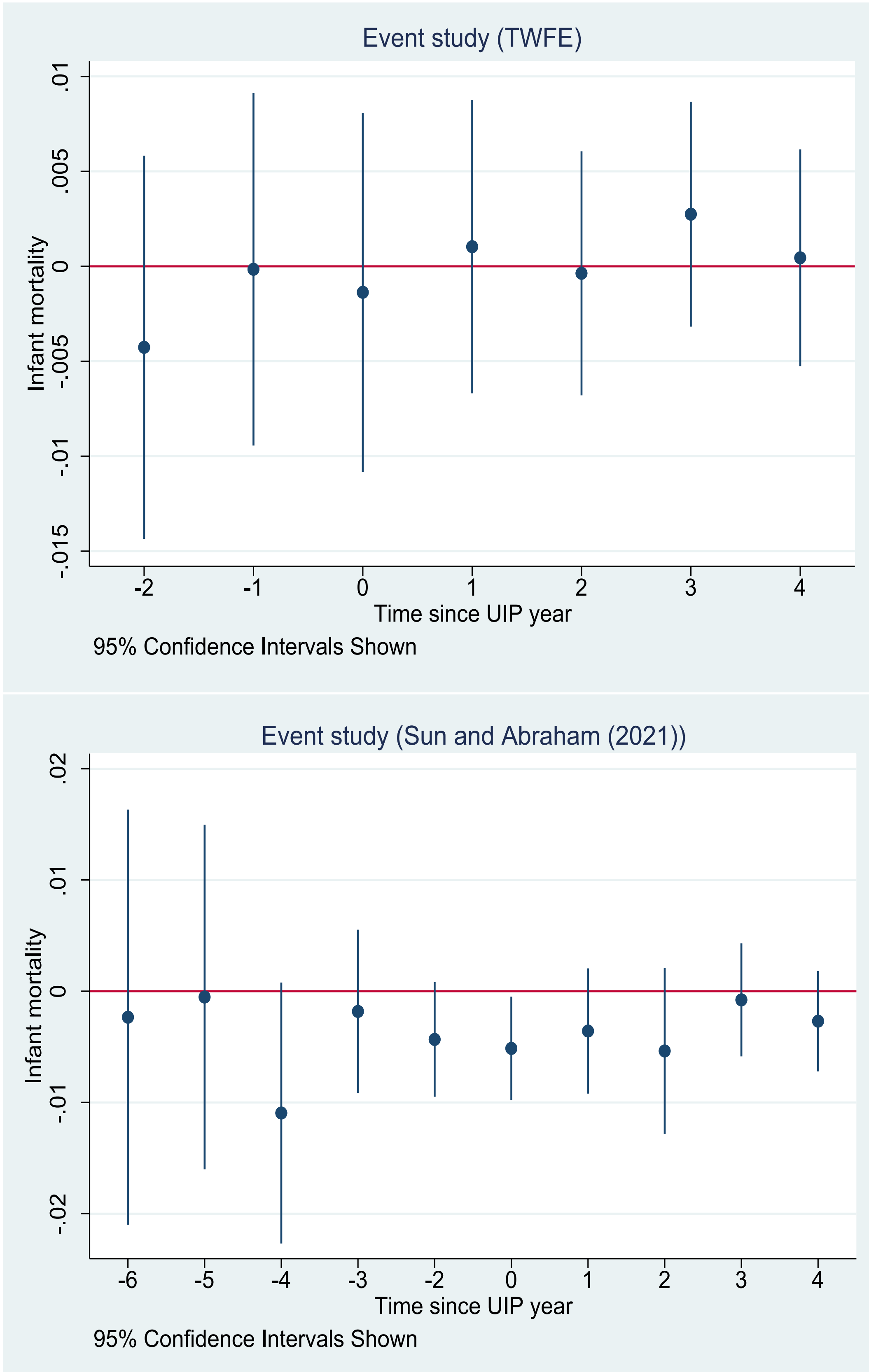


Table 4: Program Implementation

Independent Variables	UIP Launch Year (1)
Panel A: Regression based on National Rank	
National Rank in terms of child mortality before. the launch of UIP	-0.0008
	. (0.0007)
State Fixed Effects	✓
Observation	541
R Square	0.05
Panel B: Regression based on District Rank	
District Rank in terms. of child mortality before. the launch of UIP	-0.0009
	(0.0005)
State Fixed Effects	✓
N	541
R Square	0.05

Notes: Column is from estimating an ordered logit regression.Robust standard errors clustered at state level are in parentheses. National rank is in ascending order in terms of child mortality. District rank is rank of districts in terms of child mortality within a state.

Table 5: UIP effects: State-specific time trend

Explanatory Variables	Infant mortality (1)	Under-five mortality (2)
Exposed to UIP	-0.003	-0.004
	(0.002)	(0.002)
District fixed effects	✓	✓
Year of birth fixed effects	✓	✓
State specific time trend	✓	✓
N	297,385	296,511
R Square	0.04	0.06

Notes: Each column is from estimating a separate linear probability model. Robust standard errors clustered at the district level are in parentheses. *Controls*: gender and birth order of the child, mother's education, mother's age birth, religion, caste, wealth index, and rural. Survey year dummy was used. District sample weights were applied.

Table 6: Effects of UIP on Education Outcomes (state specific time trend)

Independent Variables	Literate (1)	Years of Schooling (2)	Primary school Completion (3)	Middle school Completion (4)	Secondary school Completion (5)
Exposed to UIP	0.002	0.009	-0.009***	-0.004	0.013**
	(0.003)	(0.021)	(0.003)	(0.006)	(0.006)
Controls	✓	✓	✓	✓	✓
District fixed effects	✓	✓	✓	✓	✓
Year of birth fixed effects	✓	✓	✓	✓	✓
State specific time trend	✓	✓	✓	✓	✓
N	1099940	898789	898789	898789	898789
R-squared	0.18	0.65	0.55	0.47	0.34

Notes: Columns (1), (3), (4), and (5) are from estimating a linear probability model. Columns (2), (3), (4), and (5) are from using Heckman Two-Step Method to correct for selectivity-bias. Robust standard errors clustered at the district level are in parentheses. Poor is a dummy indicating household with low socio-economic status. Survey year dummy used. Scheduled Caste(SC) and Scheduled Tribe(ST)are traditionally disadvantaged minority group. OBC is Other Backward Caste. RCH district sample weights were applied. Primary School Completion is years of schooling $\geq$5, Middle School Completion is $\geq$8 and Secondary School Completion is $\geq$10 All regressions include district and year of birth fixed effects.

Controls: gender and age of the child, caste, religion, wealth group, rural.

* shows significance at 10-percent level, ** at 5-percent level and *** at 1-percent level.