Child Cognition Benefits of the Clean India Mission Sanitation Program

Abhradeep Karmakar* Kira Villa†

January 10, 2025

Abstract

This paper examines the impact of the Clean India Mission sanitation program on children's cognition outcomes. Using survey data from the Annual Status of Education Report, and program coverage data from the Indian government, we utilize the time-varying district level roll-out of the program to examine changes in children's math and literacy test-scores. We find positive and significant improvements in math scores. This finding is robust to two different staggered difference-in-differences strategies. Literacy score effects are mostly positive, but many of these effects are statistically similar to zero. To support parallel trends, we examine changes in differential pre-trends for an unbiased treatment estimate. Multiple pathways, through which the program can affect cognition, are recognized. To address non-random program roll-out, we test if pre-treatment district-level characteristics predict program coverage. Changes in test score measurements and the respective treatment variables lead to statistically similar causal estimates.

Key words: Sanitation; Education; Clean India Mission
JEL codes: I21 Analysis of Education; I25 Education and Economic Development; O12 Microeconomic Analyses of Economic Development

*Corresponding author: Department of Economics, St. Lawrence University. Email - akarmakar@stlawu.edu
†Department of Economics, University of New Mexico. Email - kmvilla@unm.edu

1 Introduction

Open defecation practices and the lack of sanitation access are routinely linked with environmental pollution, adverse health outcomes, child mortality and low nutrition, poor living conditions, and poverty (Cutler et al., 2006; Njuguna, 2016; Mansuri et al., 2018; Spears, 2013; Rahman et al., 2020; Fink et al., 2011). Therefore, policies and programs aimed at improving sanitation access and reducing open defecation can also play important roles in improving the human capital outcomes of large groups of people, particularly in developing countries (Luby, 2014; Schmidt, 2014; Buttenheim, 2008; Patil et al., 2014; Pakhtigian et al., 2022; Coffey et al., 2018; Cameron et al., 2013; Fotio and Nguea, 2022). While targeting health or the incidence of diseases, these programs may hold added benefits to other aspects of child human capital, such as cognition. Therefore, focusing only on their health effects may undervalue the full benefit of these interventions (Santos et al., 2008; Vikram and Chindarkar, 2020; Villa, 2017).

In this paper, we evaluate the impact of the Clean India Mission (*Swachh Bharat Mission*, 2014) on children's performance on math and literacy achievement tests. This program was implemented by the Indian Government in 2014 with the aim of reducing open defecation, improving toilet-access, household/community sanitation infrastructure, and overall living conditions across the country. Attributes of the program include the construction of household and community-owned toilets, sewerage facilities, and community waste collection systems, as well as the elimination of manual scavenging (cleaning and disposal of waste from sewage pits).

Inadequate sanitation infrastructure and the associated pollution can impact children's cognitive performance through multiple potential channels. For example, Sclar et al. (2017), and Strauss and Thomas (1998), point out that poor access to proper sanitation leads to a higher risk of waste exposure and lowers living conditions. This can subsequently increase illness incidence, which can reduce school participation and cognitive outcomes. Additionally, adverse health outcomes (e.g., stunting, wasting, and diarrhea), which are associated with poor sanitation infrastructure and open defecation, are also associated with poor cognitive outcomes (Berkman et al., 2002; Fischer Walker et al., 2012; Gilmartin and Petri Jr, 2015; Petri et al., 2008). Therefore, despite not specifically tar-

geting cognition or education, household sanitation programs, such as the Clean India Mission, can impact children's cognitive performance by lowering the incidence of diseases, increasing school participation, and improving overall household quality.

To estimate the relationship between the Clean India Mission and children's math and literacy scores, we combine data from the Annual Status of Education Report with Ministry data on district-level program coverage. Due to the staggered nature of the program's implementation across districts, we estimate the group-time intent-to-treat effects of the program using the doubly robust staggered difference-in-differences strategy from Callaway and Sant'Anna (2021). Intent to treat estimates based on a not-yet-treated control group reveal that the program leads to dynamic improvements in children's math scores. These improvements begin appearing after two to four years of program exposure. Sub-group analyses reveal that both boys and girls experience positive and significant cognition score improvements due to the program. The program is also more successful in improving the cognition outcomes of children aged between 5 and 10 years. To support the findings, we further adopt the honest difference-in-differences approach developed in Rambachan and Roth (2023). Through this approach, we account for changes in pre-existing differential trends that potentially lead to biased treatment effects of the program on children's cognition.

The results are robust to multiple checks. As a test for potential mechanisms, we find that the program leads to positive changes in children's school enrollment, households' toilet ownership, and a decline in the probability of families living in a thatched dwelling. We quantify the program impacts by replacing the math and literacy scores with raw, binary, and and age-standardized test scores from the ASER surveys. We address concerns with non-random program placement by showing that the district-level coverage, and timing of implementation of the Clean India Mission is not associated with multiple pre-treatment district-level characteristics. Additional evidence of the dynamic cognition impacts of the program are reported using the methodology proposed in Wooldridge (2021). From this method, although we find a positive and significant math score impact, the average literacy score effect is statistically similar to zero. Lastly, we test if the findings

are robust to changes in the treatment variable. For this purpose, we show that an increase in the number of toilets constructed through the program per one thousand district population leads to significant improvements in children's math scores.

Literature studying the effect of household, community, and school infrastructure on improved health and better child education outcomes (e.g., school enrollment, attendance, participation, and exam-scores) is huge in scope (Alsan and Goldin, 2019; Cuesta, 2007; Jalan and Ravallion, 2003; Adukia et al., 2020; Chaudhury et al., 2006; Cuesta et al., 2016; Drèze and Kingdon, 2001; Lokshin and Yemtsov, 2005; Paxson and Schady, 2002; Schady and Paxson, 1999). By focusing on this broad area of literature, our findings contribute to the overall understanding of sanitation access in less developed countries in multiple ways. We show that children's learning achievements substantially depend on their access to sanitation facilities. This is a salient contribution of this work. This is because, while there exists vast evidence on the relationship between school and community infrastructure and children's health and nutrition, the effect of sanitation interventions on cognition is a growing area of study (Ortiz-Correa et al., 2016; Adukia, 2017; Spears and Lamba, 2016; Sharma et al., 2024; Scriptore Moreira et al., 2024; Dibaba et al., 2024). Secondly, our findings add to an even smaller group of research that focus on children's human capital benefits due to the Clean India Mission. To our knowledge, this study is the first to evaluate the program's impact on children's cognition. Thirdly, we adopt both existing and recently devised econometric methods to examine the relationship. In doing so, we exploit the staggered roll-out of the program and show that the cognition benefits sustain in later stages of exposure, in addition to an instantaneous impact.

2 Background

2.1 Clean India Mission (2014)

India has enacted several nation-wide policies to address the lack of access to basic sanitation facilities. These policies include the Central Rural Sanitation Program (1986), the Total Sanitation

Campaign (1999), the Jawaharlal Nehru National Urban Renewal Mission (2005), and the Nirmal Bharat Abhiyan (2012). Despite a heavy emphasis on eradicating open defecation, improving sanitation access and water supply, public health, waste management, and household living conditions, evidence suggests that the human capital effects of these national programs are mixed (Ghosh and Cairncross, 2014; Jha, 2010; Hueso and Bell, 2013; Elledge and McClatchey, 2013; Routray et al., 2017; Spears, 2012; Augsburg et al., 2023).

The Clean India Mission differs from existing initiatives by prioritizing positive changes in people's sanitation practices in addition to constructing household toilets in India. To improve people's sanitation practices, the program deploys social health activists, self-help groups and community organizations for establishing direct communication with households for toilet construction, organizes campaigns and social events with sports and television icons to spread sanitation and hygiene awareness, and routinely monitors waste management and disposal facilities in neighborhoods.

The program was launched in October 2014 by the newly elected government, and it was introduced in two phases. The objectives of Phase 1, implemented between 2014 and 2019, include eliminating open defecation practices in rural areas, constructing proper household toilets especially in villages, establishing mechanisms for monitoring toilet construction and community toilet usage, managing waste in communities through constructing proper sewerage facilities, eliminating manual scavenging, and maintaining adequate public waste collection systems. Between 2014 and 2019, approximately 100 million toilets were constructed and there was a remarkable decline in the percentage of Indian population practicing open defecation from 34 percent in 2014 to 19 percent in 2019 (Alexander et al., 2016). Our analysis focuses on this first phase of the program.

The broader goal of the Clean India Mission is to improve overall living conditions in communities and households, by reducing open defecation practices, and enhancing people's access to better sanitation infrastructure. This is achieved through constructing shared/community toilets and providing a cash subsidy of ₹12,000 ($144) to households for constructing their own toilets. Phase 2 of the program, occurring between 2020 and 2025, focuses on managing plastic waste, improving the lives of sanitation workers, and integrating different governmental schemes to further

enhance sanitation infrastructure.[1]

Implementation of the Clean India Mission is monitored by two different government agencies – the Ministry of Drinking Water and Sanitation in rural areas, and the Ministry of Housing and Urban Affairs in urban areas. In this paper, we study the impact of the program in rural areas, during the first phase of its coverage (i.e., from 2014 to 2018). We focus on rural areas because lack of sanitation access and open defecation practices are common occurrences in these regions. The 2019 report on Clean India Mission states that during the first phase of the program around 100 million household latrines were constructed with a 60 percent increase in the number of households having their own toilets.[2] This resulted in large improvements in people's sanitation and hygiene behaviors in rural neighborhoods and has positively affected household infrastructure and quality (UNICEF et al., 2018).

Existing research on the Clean India Mission is growing but is mostly directed towards the issues and challenges related to its implementation (Aquino et al., 2021; Bhattacharya et al., 2018; Curtis, 2019; Hutton et al., 2020; Mohapatra, 2019; Thakur et al., 2018; Kumar, 2017). A small selection of studies focuses on the effect of the program on different factors. Dandabathula et al. (2019) show that the program is associated with reduced diarrheal outbreaks during the summer-monsoon months of May, June, July, and August; both Khandelwal et al. (2020) and Hossain et al. (2022) report that the program has caused reduction in violence against women; Andres et al. (2020) find positive changes in sanitation behavior due to various campaigns related to the Clean India Mission. Similarly, research like Hutton et al. (2020) and Banerjee et al. (2017) show that the program has successfully reduced open defecation practices, improved community and household cleanliness and living conditions, improved overall household infrastructure and monetary value, and reduced pollution exposure. In a recent study, Chakrabarti et al. (2024) report correlational evidence that toilet construction through the Clean India Mission is associated with preventing 60,000 to 70,000 infant deaths annually in India.

[1] *Swachh Bharat Mission* - Phase 2 Operational Guidelines.
[2] *Swachh Survekshan Grameen* Report.

2.2 Literature review

Broad evidence demonstrates that sanitation infrastructure improves child health and nutritional outcomes like morbidity, mortality, weight-height-z-scores, and height-age-z-scores (Augsburg and Rodriguez-Lesmes, 2018; Headey and Palloni, 2019; Dickinson et al., 2015; Kumar and Vollmer, 2013; Mara et al., 2010; Pickering et al., 2015; Nandi et al., 2017; Spears, 2013; Bekele et al., 2020). There also exists a well-established body of literature that identifies factors like waste exposure, air pollution, and lead, arsenic, and toxin exposure through contaminated water, as direct outcomes of water pollution and inadequate sanitation access (Aizer et al., 2018; Bharadwaj et al., 2017; Dix-Cooper et al., 2012; Liu and Lewis, 2014; Rosado et al., 2007; Ruff et al., 1996; Saenz et al., 2018). This can subsequently result in declines in children's cognitive ability.

Some studies examine the relationship between household sanitation interventions and children's education outcomes. For example, Dearden et al. (2017), and Orgill-Meyer and Pattanayak (2020) show that children's access to better water and sanitation facilities at early ages cause improvements in their long-term cognitive outcomes. Spears and Lamba (2016) evaluate the effect of the Total Sanitation Campaign on children's cognition in India. They find that children who are exposed to the toilets constructed through the Total Sanitation Campaign by their first year of life are more likely to recognize letters and numbers at 6 years of age. Along the same lines, Cameron et al. (2021), Cronk et al. (2015), Greenberg et al. (1999), Ortiz-Correa et al. (2016), Adukia (2017) and Grantham-McGregor et al. (1999), report associations between cognitive ability or school attendance, and sanitation facilities in children's neighborhood or schools. Their findings discuss that in areas where communities lack proper sanitation infrastructure, children are more prone to waterborne diseases that affect their academic performance and school attendance negatively. Therefore, improved sanitation access can cause improvements in children's education outcomes through increased school attendance and improved overall quality of housing and community infrastructure. Early childhood exposure to quality public health services (including sanitation facilities) can also improve later life cognition, health, or employment outcomes, (Barham, 2012; Vogl, 2014; Arnold et al., 2013; Momberg et al., 2022; Null et al., 2018; Stewart et al., 2018; Tofail et al., 2018). This

suggests that programs aimed at improving sanitation infrastructure like toilets and piped drinking water can not only positively affect children in the short-run, but also yield long-run benefits.

Research like Birdthistle et al. (2011), Fentiman et al. (1999), Mathew et al. (2009), Oster and Thornton (2011), Dreibelbis et al. (2014), Freeman et al. (2014, 2012), Barrett et al. (2019), Berner (1993), Belmonte et al. (2017), Branham (2004), Durán-Narucki (2008), Flutter (2006), Hathaway (1995), McGowen (2007), Parnwell (2015), study school-level infrastructural development programs (e.g., school latrines and sanitation services, building conditions, lighting, and waste disposal), and report higher school attendance, enrollment, and completion due to these programs. However, well-targeted school-/household-level interventions may not generate anticipated effects in a predetermined area of child human capital. For example, Shah and Steinberg (2019) and Bhat (2017) find reduced test-scores due to the 2009 Right to Education Act in India, but Karmakar and Villa (2022) report improved health outcomes of children who are exposed to the mandates of the same act. Based on this argument, the Clean India Mission sanitation program is expected to have unintended positive implications for child cognition, especially in rural areas. To our knowledge, this paper is the first to examine these cognition effects of the program.

3 Data

We obtain child data from the Annual Status of Education Report (ASER), which is a repeated cross-sectional, rural household survey on education. The ASER surveys cover approximately 600,000 children aged 3 to 16 years, across 20 randomly selected rural households in 20 to 30 villages in around 600 rural districts. The surveys collect information on household size, parental education, household assets and infrastructure, gender, and schooling status of children. It also tests all children aged 5 to16 years in basic literacy (reading) and math proficiency. We use data from six ASER rounds executed in 2011, 2012, 2013, 2014, 2016, and 2018.

The ASER literacy and math achievement tests score learning proficiency on a scale from 0 to 4. The literacy assessment has five levels: cannot read (0), can read letters (1), can read words

(2), read grade 1 text (3), and read grade 2 text (4). The five assessment levels for math include cannot do basic arithmetic (0), can recognize single-digit numbers (2), can recognize double-digit numbers (3), can perform two-digit subtraction (3), and can perform three-by-one digit division (4). Figure 1 presents the distribution of raw literacy and math scores for the sample children. Averaged across all ages and survey, just over half the sample falls in the bottom three math scores, meaning that they either cannot do basic arithmetic or can only recognize up to one- or -two-digit numbers. Only approximately a quarter can perform a three-by-one-digit division. Regarding literacy, approximately 43 percent can read at least grade 2 texts but 30 percent of the sample either cannot read at all or only recognize letters.[3]

Data on the annual coverage of the Clean India Mission come from the Indian government's Ministry of Drinking Water and Sanitation. Coverage of the program in each district or block (smaller geographical regions than districts) is reported as the number of households that had a toilet constructed under the program annually from 2014 through 2021. Using these data, Figure 2 illustrates the percentage of households in each district that are covered by the Clean India Mission from 2014 to 2018. In 2014, 2015, and 2016, approximately 25 percent of households, on average, across districts are covered by the program. This average increases to around 50 percent in 2017, and then to over 75 percent in 2018. To put it together, Figure 2 shows a gradual and remarkable increase in household toilet coverage through the program from 2014 to 2018.

We match the district-level program coverage data for 2014, 2015, and 2017, with the ASER rounds of 2014, 2016, and 2018. Using the previous year's coverage ensures that children experienced one full year of program exposure. The repeated cross-sectional ASER surveys, upon merging with the coverage data from the Ministry of Drinking Water and Sanitation, yield a total of 491 districts in each survey round. Among these, 1,743,110 children (in 460 districts) reside in districts first covered by the program by the 2014 ASER round, 80,357 children (in 27 districts), reside in districts that are first treated by ASER 2016, and 13,267 children (in 4 new districts) reside in districts first exposed to the program by ASER 2018. This means that we do not have a

[3]In all the tables and remaining figures, math stands for math scores and literacy stands for children's reading scores.

never-treated district in the study, as by ASER 2018 all of the 491 sample districts are covered by the Clean India Mission.[4]

Trends in standardized math and literacy scores, along with the percentage of households with a toilet are shown Figure 3. These trends are shown separately for districts first covered by the program by the 2014 ASER round (in blue), 2016 ASER round (in green) and the 2018 ASER round (in red), respectively. In panels A and B, we observe a downward trend in math and literacy scores 2013 in all the three treatment groups. After 2013, this downward trend levels off and even reverses slightly in districts first covered by the Clean India Mission in 2014 and 2016. In fact for the 2014 treated districts, some positive changes begin to appear from 2016. These changes are even larger in 2018. For 2016 treated districts, small math and literacy score improvements can be noticed from the 2014 survey year. In years following the implementation of the Clean India Mission, average math and literacy scores in districts treated in 2014 and 2016 are larger than those in the 2018 treated districts. These math scores differences begin to appear in 2016, and literacy scores differences can be seen from 2014, respectively.[5]

Panel C of Figure 3 shows large differences in the percentage of households with a toilet between districts first treated in 2014, and those treated in 2018 or 2016. These differences become smaller as we move from 2011 to 2018. There also appears to be a clear increase in the percentage of households' toilet ownership in years following the implementation of the Clean India Mission. These years are 2014, 2016, and 2018, respectively. We find that in 2011, approximately 40 percent, 75 percent, and 60 percent of households have their own toilet in the 2014, 2016, and 2018 treated districts. In 2018, which represents four years after the implementation of the program, these percentages increase sharply to 70 percent in 2014 treated districts, 90 percent in 2016 treated districts, and 80 percent in 2018 treated districts, respectively.

Table 1 reports summary statistics for baseline characteristics (2011) for each treatment group.

[4]Appendix Figure A.1 shows the districts first covered by the program in 2014 (treatment group in 2014), covered by the program in 2015 (treatment group in 2016), and covered by the program in 2017 (treatment group in 2018) in panel A. Panel B shows the same groups for the 491 districts used in this analysis.

[5]Raw scores are standardized using the district-specific means and standard deviations over the entire duration of the ASER surveys (2011 to 2018). The standardized math and literacy scores represent the standard deviation (SD) differences between the raw scores of a child and the mean of raw scores in the district.

Columns 1, 2, and 3 report the average characteristics for districts first treated in 2014, 2016, and 2018, respectively. Differences in these average characteristics between each of the three treatment groups are reported in columns 4, 5, and 6. Standardized math and literacy scores are statistically similar between children in districts treated in 2014 and 2016. But we find that children in districts treated in 2018 report statistically larger math and literacy scores than those in districts treated in 2016 or 2014. The sample is adequately represented by both boys and girls, who report a mean age of roughly 10.5 years.[6]

Overall, there appears to be non-random program placement in that the earliest-treated districts (those treated in 2014) are worse off in numerous dimensions than districts treated later. Children in 2014 districts are slightly less likely to have parents who have attended a school. They are also less likely to live in districts with access to village-level facilities like post-offices, and health centers than districts first treated in 2016 or 2018. Differences in the proportion of villages with a road are statistically significant, however these differences are very small with approximately 75 percent of villages having road access regardless of treatment year. Households in the 2014 treated districts are significantly less likely to own a TV, mobile phone or have access to electricity than the 2016 and 2018 treated districts.

When we compare the 2016 and 2018 treated districts, these differences appear to be less clear and small in magnitude. Table 1 also shows that approximately 35 percent of the sample lives in a thatched house in 2014 and 2018. This estimate is around 19 percent on average for the 2016 treated districts. Approximately 95 percent of the sample children are enrolled in school, with approximately 70, 61, and 84 percent of children enrolled in government school in the districts first treated by 2014, 2016, and 2018.

Using the district-level population data from the 2011 population census, we convert the total number of toilets constructed through the Clean India Mission by 2018, into the program coverage rate per one thousand district population. The bottom row in Table 1 reports this estimate. We ob-

[6]Figure 4 plots the age distribution of the sample children in each of the three treatment groups. These distributions show a high representation of children aged around 8, 10, and 12 years. For 2018 treated districts, the distribution is more skewed towards younger children than the 2016 and 2014 treated districts.

serve that by 2018 approximately 53.68, 33.91, and 10.21 toilets are constructed per one-thousand people in the 2014, 2016, and 2018 treated districts, respectively. Thus, intuitively, longer exposure to the program resulted in more toilets constructed per one-thousand population by 2018.

4 Empirical strategy

Due to differences in treatment timing, results from a standard two-way-fixed-effects strategy may be biased because earlier treated units are used as controls against the future treated units (Borusyak et al., 2021; Roth, 2022; De Chaisemartin and d'Haultfoeuille, 2020; Sun and Abraham, 2021; Athey et al., 2016; Goodman-Bacon, 2021; Sant'Anna and Zhao, 2020; Wooldridge, 2023). To address the staggered nature of program implementation by which 460 districts are covered by 2014, 27 districts are covered by 2016, and 4 districts are covered by 2018, we utilize the doubly robust staggered difference-in-differences approach proposed by Callaway and Sant'Anna (2021). Using their method, we estimate the group-time treatment effects and the aggregated treatment effects of the Clean India Mission on children's cognition. Then we estimate the heterogeneity in program impact across children's age and gender. These results are shown in section 4.3. Section 4.4 tests for pre-treatment trends and reports the aggregated event study results. In section 5, we present multiple robustness checks. To address the violation of parallel pre-treatment trends, we adopt the honest difference-in-differences approach from Rambachan and Roth (2023). We identify multiple mechanisms through which the program can affect children's cognition. We examine program impacts by replacing the standardized outcome variables with binary and raw math and literacy scores. To address concerns with non-random program roll-out, we test for the association between pre-treatment district-level characteristics, and the program's district-level coverage and timing of coverage. To support the difference-in-differences results, we estimate dynamic treatment effects of the program using the method from Wooldridge (2021). Lastly, we use the continuous program coverage as the explanatory variable and measure the effects of the number of toilets constructed per one thousand district population on children's math and literacy

scores.

4.1 Staggered difference-in-differences

Following Callaway and Sant'Anna (2021), the group-time or district-time intent to treat effects (ITT) of the Clean India Mission using a staggered doubly robust estimator are formalized using equation (1):

$$ITT(g,t) = \mathbb{E}\left[\left(\frac{G_{ig}}{\mathbb{E}\left[G_{ig}\right]} - \frac{\frac{P_t\left(X_{id}\right).C_i}{1-P_t(X_{id})}}{\mathbb{E}\left[\frac{P_t(X_{id}).C_i}{1-P_t(X_{id})}\right]}\right).\left(Z_{it} - Z_{i(g-1)} - OR^C\right)\right] \tag{1}$$

where, for a child i in district d, G_{ig} is a binary variable indicating that the district is first covered by the Clean India Mission in year g. These years are 2014 and 2016, respectively. C_i represents the control group, which comprises the not-yet treated districts. C_i also includes districts treated in 2018. $P_t(X_{id})$ and $1 - P_t(X_{id})$ are the propensity scores used for weighting the control group. These propensity scores represent the predicted probabilities of being or not-being treated in year g. $\frac{P_t\left(X_{id}\right).C_i}{1-P_t(X_{id})}$ is the ratio of the predicted probabilities of being or not-being treated in year g. By multiplying this ratio with the binary indicator for the control group, we put more weight on children in the not-yet treated districts (and 2018 treated districts) that have similar district level characteristics as those first treated in year g, and less weight on children in the not-yet treated districts (and 2018 treated districts) whose characteristics differ from those first treated in year g. This reweighting procedure ensures that the common support assumption is satisfied, and that the control group distribution is similar to the treatment group, in terms of average characteristics.

Z_{it} and $Z_{i(g-1)}$ are the standardized math and literacy scores for children in the respective treatment district at year t, and the year immediately before treatment. This year is represented

as $g-1$. OR^C is the outcome regression notation, which shows the conditional mean of the differences in standardized math and literacy scores between years t and $g-1$ for children in not-yet treated districts and 2018 treated districts. It is represented as: $\mathbb{E}[Z_{it} - Z_{i(g-1)}|X_{id}, C_i]$.

Intuitively, $ITT(g,t)$ is the group-time or district-time intent to treat effect representing the difference between the expected change in cognition scores of children in treated districts (2014 and 2016), between periods t and $g-1$, and the expected cognition score changes of children in control districts, between periods t and $g-1$. t represents each survey year following the first year of program coverage (g). For districts treated in 2014, t equals 2014, 2016, and 2018. For districts treated in 2016, t equals 2016, and 2018.

Calculation of $ITT(g,t)$ under the staggered doubly robust approach relies on the outcome regression and inverse probability weighting methods. Outcome regression utilizes the conditional expectation of the outcome evolution for the not-yet treated and 2018 treated districts (Heckman et al., 1998, 1997). Inverse probability weighting utilizes the conditional probability of being treated in a given period Abadie (2005).

The $ITT(g,t)'s$ are conditioned on several covariates (X_{id}) which include whether the child's parents have attended school, child's gender, age, and the quadratic of age. These covariates also include district level information from the 2011 ASER survey, which are the percent of villages with a road, a post office, a healthcare center, percent of school enrolled children, percent of children in government schools, percent of households with a TV, with electricity, and the percentage of households living in a thatched dwelling.

From all the group-time or district-time intent to treat estimates, we quantify the weighted average of the program's impact on children's math and literacy scores. We report the single aggregated and partially aggregated parameters of the group/district-specific estimates, event-time effects and the calendar/survey-time estimates. We also estimate the simple weighted average of the ITT for boys and girls, and for children in different age-groups.

4.2 Threats to identification

The primary assumption required for difference-in-differences models is that of parallel trends. A violation of this assumption means that even without the program, children's test scores in treatment districts were trending differently than those in control districts, and thus differences in post-treatment cognition changes between the groups cannot be entirely attributed to the program. We address this concern by testing for pre-treatment parallel trends using an aggregated event study design. Failing this test indicates that the difference-in-differences findings may simply be capturing pre-existing differences in trends between the treatment and control groups. Because we find evidence for differences in pre-treatment trends, we check the sensitivity of the estimated intent to treat effect to this violation using the honest difference-in-differences approach from Rambachan and Roth (2023).

To support the hypothesis that the test score improvements are indeed driven by the Clean India Mission, we test for multiple potential mechanisms through which the program plausibly affects children's test outcomes. At the child-/household-level, these mechanisms include whether household has its own toilet, whether child is enrolled in a school, and whether the household lives in a thatched dwelling. We also examine if the findings are sensitive to how the test score variable is defined. Given the categorical measurements (0, 1, 2, 3, and 4) of test scores, children's math or literacy scores may not differ from the sample mean by large amounts. To address this concern, we estimate our model using the raw, binary, and age-standardized math and literacy scores as the dependent variables.

The validity of our research design also requires the exogeneity of the roll-out of the Clean India Misson (i.e., there is not a non-random program placement). This means that the timing of program roll-out is not determined by district-level characteristics. We address this concern by checking if pre-treatment district-level attributes predict toilet construction through the program, and the timing of program implementation. If we find that these characteristics significantly affect program roll-out, then the difference-in-differences results are likely biased due to non-random program placement.

To support the difference-in-differences findings, we adopt the staggered approach from Wooldridge (2021) and measure the dynamic treatment effect of the Clean India Mission on children's cognition. This model relies on all possible interactions between treatment timings and survey times. Lastly, because the results are based on a differential timing in the district-level program implementation, we are unable to examine the cognition differences between children in districts with different levels of program coverage. To address this concern, we use a continuous explanatory variable and examine the effects of the number of toilets constructed through the Clean India Mission on children's cognition. All of these checks largely support the main findings and are discussed in further detail in Section 5.

4.3 Results

First, in Table 2, we report the aggregated treatment effect estimated by equation (1). Panel A reports the weighted average by group size of all the group-time or district-time intent to treat effects. Coefficients on $Weighted\ average\ (ITT)$ indicate that children in treatment districts report significantly larger improvements in their math and literacy scores by 0.166 standard deviations (SD) and 0.195 SD, than those in the control districts. In Panel B, the coefficients on $Treated\ in\ 2014$ and $Treated\ in\ 2016$ represent the group-specific or district-specific effects of the program. The average math score effects for children in both the 2014 and 2016 treated districts indicate that toilet construction through the Clean India Mission causes significant improvements in children's math scores regardless of the first treatment year. The effect on literacy scores is only positive and significant for children in the 2014 treated districts. This effect is around 0.19 SD. Panel C shows the program effects by each post-treatment survey year. The effect on math scores is positive but it is only significant in 2018. In addition to this, we find that the math score impact increases with longer program exposure, as the estimate in 2018 (0.491 SD) is larger than that in 2016 and 2014. The literacy score effects in 2014, 2016, and 2018 are positive, statistically significant, and also increasing over time. Across all three panels, coefficients on $Weighted\ average\ (ITT)$ and $Single\ parameter$ are nearly similar. These effects are around 0.16 SD for math scores, and 0.19

SD for literacy scores.

Following this, we examine if the cognition benefits of the Clean India Mission differ between boys and girls, and younger and older children, given existing evidence that cognitive production or human capital investment behavior varies across these groups (Muralidharan and Sheth, 2016; Grant and Behrman, 2010; Jayachandran, 2015; Glick, 2008; Patrinos, 2000; Gevrek et al., 2018; Hermann and Kopasz, 2021; Machin and Pekkarinen, 2008; Lavy, 2015). Table 3 reports the weighted average of all intent to treat estimates for boys and girls in Panel A. The impact of Clean India Mission on both math and literacy scores are positive and statistically significant. Boys in treated districts report significantly larger improvements in math and literacy scores by 0.151 SD and 0.198 SD, respectively. For girls, these improvements lie around 0.178 SD and 0.185 SD. When we split the sample in two different age-groups, which are 5 to 10 years and 11 to 16 years, we find that toilet construction through the Clean India Mission is more successful in improving the cognition outcomes of children aged between 5 and 10 years. These aggregated improvements are around 0.196 SD and 0.258 SD, respectively. The estimated effects on older children are small in magnitude and those are accompanied by large standard errors.

To examine the dynamic effects of the Clean India Mission, we report group-time or district-time intent to treat effects on children's math and literacy scores for districts treated in 2014 and 2016 in Table 4. The corresponding estimates are reported in Figure 5. Pre-treatment estimates are shown by the blue bars, and the post-treatment program impacts are shown by red bars. In both these models, the control group is composed of districts that are not-yet-treated. The comparison year for the post-treatment estimates is the year immediately prior to treatment. These years are 2013 in Panel A, and 2016 in Panel B. Comparison year for each pre-treatment year is the respective preceding year.[7]

We find positive math score effects associated with the Clean India Mission that increase the longer districts are exposed to the program. By two years post-program implementation, the program's impact on math scores in 2014 treated districts is small in magnitude and statistically in-

[7]Results without covariates can be found in the Appendix in Tables A.1 and A.2, and Figure A.2.

significant. However, by 4 years post-implementation, math scores in 2014 treated districts are approximately a half a standard deviation (SD) higher, on average (Figure 5, panel A). For districts first treated in 2016, we find that the program results in an average increase in math scores of approximately 0.17 SD two years after implementation (Figure 5, panel C). The fact that the 2016 treatment group experiences significant improvements in math scores sooner than 2014 treated districts may be due to systematic differences between the two treatment groups or due to learning on the part of those implementing the program (i.e., execution of the program improves over time). Therefore, the post-treatment coefficients point to a dynamic effect of the Clean India Mission on children's math scores in districts that are first covered by the program in both 2014 and 2016. These dynamic improvements begin appearing within two to four years of continuous program exposure. Point estimates on period zero that represent the instantaneous program effects are less clear and statistically similar to zero. Turning to literacy scores, like with math scores, we estimate a significant and positive effect of the program that improves over time for districts first treated in 2014 (Figure 5, panel B). According to these results the program is associated with a 0.06 SD average increase in literacy scores in the year of implementation that increases to 0.18 SD two years after implementation and to 0.34 SD four years after implementation. Conversely, we estimate a null effect on literacy scores for districts first treated in 2016 (Figure 5, panel D).

In Figure 5 (panel A), we find no evidence of a violation of parallel pre-treatment trends in math scores for the 2014 treatment group; estimated effects for this group prior to treatment (the blue bars) are all statistically insignificant and close to zero. However, we do find a violation of parallel pre-treatment trends in the math scores for districts first treated in 2016. Specifically, we estimate a significantly positive estimate 3 years prior to treatment and a significantly negative estimate 1 year prior to treatment. Turning to the literacy score effects in panels B and D. For both groups (those first treated in 2014 and 2016), we observe violations of parallel pre-treatment trends. The violations of parallel pre-treatment trends observed for literacy scores of the 2014 treatment group and for both math and literacy scores of the 2016 treatment group raise concerns of bias in the estimated ITT effects for these outcomes.

4.4 Aggregated event study

To further evaluate dynamic treatment effects and pre-treatment trends, we estimate an aggregated event study model. This model allows us to check for average parallel pre-treatment trends between the treatment and control groups. Event studies also help in examining the timing in appearance of treatment effects, whether the effects are increasing or decreasing over time, and whether the effects are permanent or transient (Miller et al., 2021; Clarke and Tapia-Schythe, 2021; Cunningham, 2021; Freyaldenhoven et al., 2019).

We report the aggregated event study results in Table 5. Coefficients on $-3, -2, -1, 0, +2,$ and $+4$ are the intent to treat estimates by length of program exposure. For the post-treatment measures (lag estimates), the base year is the year immediately prior to treatment. Base year for each pre-treatment (lead) estimate is the respective preceding year. A further aggregation of all lead and lag effects is represented by the coefficients on $Pre - treatment\ average$ and $Post - treatment\ average$ in the table. The corresponding intent to treat lead and lag estimates are plotted in Figure 6.[8]

In Figure 6, panel A, the coefficient for period $+4$ is positive and significant. This estimate is 0.495 SD, respectively. Coefficients on 0 and $+2$ are statistically insignificant, which indicates that it takes up to four years for the dynamic effect of the Clean India Mission on children's math scores to start showing. The post-treatment literacy estimates in panel B represent both an instantaneous and dynamic effect of the program. The instantaneous effect is reliant on the positive and significant point estimate of 0.067 SD at period zero. This effect becomes larger over time pointing to dynamic literacy score improvements in periods $+2$ and $+4$.

Similar to the group-time estimated effects reported in Figure 5, we find evidence of violations of common pre-trends in the aggregated event study results reported in Figure 6. For math scores, we have a significant pre-treatment coefficient at period -3, and for literacy scores, all the pre-treatment coefficients are statistically significant. Because of the clear violations of parallel trends, our estimates of the post-treatment dynamic math and literacy score changes may not be entirely

[8]Event study results without covariates can be found in the Appendix in Table A.3, and Figure A.3.

due to the Clean India Mission making the intent-to-treat estimates biased. This is a problem that we address in the following section.

5 Robustness

5.1 Violating parallel trends

Evidence presented in Figures 5 and 6 demonstrate a violation of the parallel assumption, which implies bias in the post-treatment estimates. To assess the extent our findings are influenced by this violation of parallel pre-treatment trends, we adopt the honest difference-in-differences approach developed in Rambachan and Roth (2023). Using their relative magnitudes method; we test for the sensitivity of the dynamic intent to treat estimates at period $+4$ in the event study analysis (in Table 5 and Figure 6), to different changes in the maximal pre-treatment trend violation. These dynamic estimates at period $+4$ are 0.495 SD and 0.346 SD for math and literacy scores, respectively.

The existence of differential pre-treatment trends implies that there may be pre-existing district-specific cultural, political, economic, and environmental factors (e.g., elections, school-level policies, government policies, natural or health emergencies) causing the test scores of children in treated districts to trend differently than those in control districts. Rambachan and Roth (2023) propose a method that extrapolates observed differential pre-treatment trends into the post-treatment period to generate a set of counter factual trends. Specifically, following their method, we introduce a relative magnitudes parameter (M). This parameter stands for the proportion of the maximal (or worst) pre-treatment differential trends, that the counterfactual post-treatment differential trends can be equivalent to in order to obtain a null effect of the program. From Table 5, the maximal (or worst) pre-treatment differential trend estimates for math and literacy scores are 0.202 SD and 0.139 SD, respectively.

The relative magnitudes approach is formalized as: $Post_{trends} = M \times Pre_{trends}$, where $Post_{trends}$ represents counterfactual post-treatment differential trends, and Pre_{trends} represents the maximal (or worst) pre-treatment differential trend estimate. M equals 1 if differential trends

are assumed to be constant over pre- and post-treatment periods (i.e., $Post_{trends} = Pre_{trends}$). Other values of M that we test are $0.5,\ 0.75,\ 1.25,\ 1.5$ and 1.75. $M = \ 1.5$ or 1.75 indicates that $Post_{trends}$ are 1.5 or 1.75 times the maximal existing Pre_{trends}. $M = \ 0.5$ indicates that $Post_{trends}$ are one-half of pre-existing differential cognition trends between children in treated and controlled districts.

From the respective relative magnitudes (M) values, we obtain a breakdown point. At this breakdown point the intent to treat estimate at period $+4$ (from Table 5 and Figure 6) is statistically similar to zero (i.e., no longer significant). This means that at the breakdown point, the differential cognition effect of district-specific pre-existing cultural, political, economic, and environmental factors, are large enough for Clean India Mission to not produce a statistically meaningful impact on children's math and literacy scores.

Using the honest difference-in-differences approach, we obtain multiple confidence intervals on the intent to treat estimate at period $+4$. These confidence intervals tell us if the program's treatment effects at period $+4$ are sensitive to relative changes in pre-treatment differential trends, between the pre- and post-treatment periods. Table 6 presents the honest difference-in-differences results. This table reports confidence intervals on the intent to treat estimate using the relative magnitudes approach. These confidence intervals are plotted in Figure 7. The red bars show the bounds on the original intent to treat estimate at period $+4$ from the event study design (in Table 5 and Figure 6), and the blue bars show the bounds on the treatment estimate for different values of M.

In Figure 7, we find that that if post-treatment differential trends are restricted to be equivalent to the maximal pre-treatment differential trend (i.e., $M = 1$), the intent to treat estimates on both math and literacy remain positive and statistically significant. The lower and upper bounds are 0.112 and 0.877 for math, and 0.106 and 0.587 for literacy, respectively. Looking further to the right, in panel A the breakdown point of M is 1.5 (Confidence intervals are -0.054 and +1.044). This indicates that the counterfactual post-treatment differential trends need to be equivalent to 1.5 times as large as the maximal (or worst) pre-treatment differential trends, for the dynamic effect

of Clean India Mission on children's math scores to be statistically similar to zero. In panel B, the respective breakdown point, at which we fail to recognize a statistically meaningful impact on children's literacy scores, is $M = 1.75$. The breakdown points on both math and literacy scores are quite high which tells us that for Clean India Mission to not produce a dynamic effect on children's cognition, the post-treatment effect of pre-existing district specific cultural, political, economic, and environmental factors need to be almost twice as much as their largest effect in the pre-treatment period.

Drawing on the group-time or district-time intent to treat effects from Table 4 and Figure 5, we test for the sensitivity of the post-treatment math and literacy score estimates for each treatment group. These post-treatment estimates are on period $+4$ for districts treated in 2014, and period $+2$ for districts treated in 2016. From Table 4, the maximal (or worst) pre-treatment differential trend for literacy score is 0.07 SD for districts treated in 2014, and 0.139 SD for districts treated in 2016. The maximal (or worst) pre-treatment differential trend for math score is 0.202 SD for 2016 treated districts. We use 0.004 SD as the maximal pre-treatment differential math score trend for 2014 treated districts, although this coefficient is statistically insignificant.

Results are shown in Table 7 and Figure 8. In panel A of Figure 8, We find that for the post-treatment math score estimate to lose statistical significance, differential pre-treatment trends have to change by a huge relative magnitude of approximately 6.5 points. For districts treated in 2016, however, counterfactual post-treatment differential trends need to be equivalent to one-half as large as the maximal (or worst) pre-treatment differential trends, for the dynamic effect of Clean India Mission on math scores to be statistically similar to zero. The literacy score breakdown point for 2014 treated districts (panel B) is 2.5 which is slightly larger than the estimated breakdown point in the honest difference-in-differences approach based on the aggregated event study results. Lastly, regardless of the M values, we find no statistically meaningful changes in the post-treatment literacy score estimates for districts treated in 2016.[9]

[9]Honest difference-in-differences results using the smoothness bounds approach are reported in Appendix Tables A.4 and A.5, and Figures A.4 and A.5.

5.2 Potential mechanisms

There are multiple channels through which better sanitation can improve child school achievement outcomes. Improved access to better sanitation facilities can reduce the likelihood of children being affected by short-term illnesses, which, in turn, can positively influence their school attendance (Battiston et al., 2013; Dreibelbis et al., 2013a,b; Rauniyar et al., 2011) and subsequent school performance. By providing grants for toilet construction, Clean India Mission can also improve the socioeconomic conditions of a household by freeing resources for other uses and improving household health and living conditions through improved community and household cleanliness, better household infrastructure and value, and reduced exposure to pollution (Hutton, 2013; Banerjee et al., 2017; Behera et al., 2021). All these factors can contribute to better school outcomes for children.

We explore these potential mechanisms for the findings by estimating whether the program leads to changes in household-level access to a toilet, improves school enrollment of children, and reduces the likelihood of household living in a thatched dwelling (a proxy for economic well-being). To do so, we use the same Callaway and Sant'Anna (2021) model we use to obtain the main results reported in Table 2.

Weighted averages of the intent to treat effects of the Clean India Mission on these outcomes are reported in Table 8. The change in the probability of households having their own toilets is larger for treated children, but the point estimate is not statistically significant. We find a small improvement in children's school enrollment due to the program but, again, the coefficient is statistically similar to zero. One possible reason for this is that more than 90 percent of the children in all three treatment groups are enrolled in school, meaning that there may be little room to impact this outcome. However, it is possible that the program improves school $attendance$, even if there is little impact on the extensive margin of school enrollment. Unfortunately, the ASER data do not contain attendance information, so we cannot check this. Finally, we find that the program leads to a significant decline in the likelihood of households living in a thatched house. Thatched dwellings are associated with lack of access to proper sanitation facilities, lower socioeconomic

status, and increased incidences of illnesses (Prithviraj et al., 2017; Azurin and Alvero, 1974). Therefore, a decline in probability of living in a thatched dwelling indicates that toilet construction through the Clean India Mission causes improvements in overall household living conditions and socioeconomic well-being, which could further transform into better cognitive achievements for children.

5.3 Test score measurement

In the main findings we use standardized test scores as the outcome variables. In Table 9, we examine the sensitivity of our findings to three alternative constructions of the test score data. Following some others who have employed the ASER test score data (Chakraborty and Jayaraman, 2019; de Barros et al., 2022; Das and Sarkhel, 2023; Bhattacharya et al., 2023), first we re-estimate the model using the raw test scores as the outcome variable; second, we re-estimate the model using binary measures of math and literacy skills as the dependent variable; and third, we re-estimate our model using math and literacy scores that are standardized using the age-specific mean and standard deviation values. Remember, the ASER math and literacy variables capture assessments of children's learning proficiency at five categorical levels (levels 0, 1, 2, 3, and 4) that are increasing in skill level (as described in Section 3).[10] We, therefore, generate binary variables for math and literacy, which are equal to one if the child tests at level four proficiency (the highest proficiency level) and zero otherwise. For math, level four indicates that the child can perform three-by-one-digit division. For literacy, level four indicates they can read grade two text. Results using the raw scores as the dependent variables are reported in columns 1 and 2 of Table 9, those using the binary measures are reported in columns 3 and 4, and those using the scores that are standardized by each age are shown in columns 5 and 6. All the results conform to the main findings and indicate that Clean India Mission results in improvements in performance in math and literacy.

[10]Literacy (reading) assessment has 5 levels: cannot read, read letters, read words, read grade 1 text, and read grade 2 text. Math assessment comprises, one cannot do basic arithmetic, can recognize single-digit numbers, can recognize double-digit numbers, can perform two-digit subtraction, and can perform three-by-one-digit division.

5.4 Determinants of program implementation

Another concern with evaluating government programs relates to the endogenous nature of the program roll-out. Table 1 indicates that the earliest treated districts were worse off in numerous dimensions. This indicates non-random program rollout, meaning that the implementation of the Clean India Mission could be driven by different household, village, or district-level factors. We address this concern examining the association between different pre-treatment district-level characteristics and the number of toilets constructed (per 1000 population) in the district through the program in its first year of implementation. We also examine the association between pre-treatment district-level characteristics and the timing of program implementation in the district. To do this, we use two different dependent variables that measure program implementation timing. The first dependent variable equals one if a district is first treated in 2014 and equals zero for districts treated in 2016 or 2018. For the second variable, we combine the 2014 and 2016 districts. This dependent variable equals one if a district is first treated in 2014 or 2016 and equals zero for districts treated in 2018.

We use the pre-treatment ASER rounds (2011, 2012, and 2013), and aggregate the data to the district level. These data include the percentage of villages in the district with a post-office, bank, and health center, the percentage of households living in a thatched dwelling, percentage of households with a TV, percentage of households with a mobile phone, and percentage of households with a toilet. We also utilize data on percentage of school enrolled children, percentage of children in a government school, and the percentage of children whose parents have attended a school.

Results are shown in Table 10. When we use the continuous coverage of the program as the dependent variable, we find that most of the estimated coefficients are statistically insignificant, except the weakly significant estimates on parental school attendance, and children's school enrollment. However, the coefficients on percentage of fathers having attended a school, and the percentage of school enrolled children are negative in magnitude. In columns 2 and 3, the coefficients on all the pre-treatment district-level characteristics are also close to zero. Except for the coefficients on percentage of fathers having attended a school, percentage of households with

electricity, and the percentage of households living in a thatched dwelling in column 1, all the estimated associations are statistically similar to zero. The lack of a methodical association between any of these characteristics and the dependent variables (level of program coverage and timing of program implementation), indicates that the district-level roll-out of the Clean India Mission is unlikely to be endogenous in nature, and our results are unlikely to represent biased treatment estimates due to a non-random program implementation.

5.5 Staggered difference-in-differences using an alternative specification

To support the main results, we examine the program effects using the staggered difference-in-differences method following Wooldridge (2021). Equation (2) presents the strategy:

$$Z_{idt} = \sigma_0 + \sum_{g}\sum_{t=g}^{2018}\sigma_{gt}.(Treat_{dgt}) + X + \theta_d + \theta_t + \zeta_{idt} \tag{2}$$

where, for a child i in district d at time t, Z_{idt} is the standardized math or literacy score. X is the same set of covariates from equations (1), and θ_d and θ_t are district and year dummies. g is the first treatment year of a district. g can be 2014, or 2016. t represents each survey year following the first year of program coverage. For districts treated in 2014, t equals 2014, 2016, and 2018. For districts treated in 2016, t equals 2016, and 2018. $Treat_{dgt}$ is a binary treatment indicator. It equals one if a district d is treated in year g, at period t, and it equals zero otherwise. The control group is composed of districts that are not-yet-treated and the four districts that are treated in 2018. σ_{gt} is the intent to treat effect, which represents the differential cognition impact of the Clean India Mission on children in districts treated in 2014 or 2016 relative to those in the control districts. These intent to treat estimates are reported for each post-treatment year, i.e. years following the first program coverage. This is a similar strategy to the one adopted in Nagengast et al. (2024).

Results are reported in Table 11 and plotted in Figure 9. By comparing the findings from Table 4 and Table 11, we find that the staggered difference-in-differences strategies using both Callaway

and Sant'Anna (2021), and Wooldridge (2021) yield dynamic math score improvements for children in the 2014 and 2016 treated districts. These improvements start appearing within two (for 2016 districts) to four (for 2014 districts) years of continuous program exposure. We also find robust dynamic literacy score improvements for the 2014 treated districts. But the estimated effects on literacy scores for districts treated in 2016 are less clear. This is because, while both the instantaneous and dynamic effects using Wooldridge (2021) are positive and statistically significant, the group-time intent to treat literacy estimates using Callaway and Sant'Anna (2021) for 2016 districts are statistically similar to zero.

In the lower half of Table 11, we report the aggregated intent to treat estimates of the program. We find that children in treatment districts report significantly larger improvements in their math scores by 0.041 SD, relative to those in the control districts. The aggregated literacy score effect is positive but not statistically significant. This estimate is 0.028 SD. Although all the aggregated estimates are positive, it is noteworthy that the aggregated effects from Wooldridge (2021) are slightly smaller than the weighted average of the intent to treat estimates from Callaway and Sant'Anna (2021) which are reported in Table 2.

5.6 Continuous program coverage

Lastly, we examine the effect of the number of toilets constructed through the Clean India Mission on children's math and literacy scores. These results are reported in Table 12, where $Coverage\ per\ 1000$ is the explanatory variable. We also include the quadratic of $Coverage\ per\ 1000$ in the regressions. $Coverage\ per\ 1000$ represents the number of toilets constructed under the program divided by the total district population, multiplied by one thousand. Coefficients on this variable retrieve the average linear effect of one additional toilet in the district on children's math and literacy scores.

We find that the construction of one toilet per thousand population in the district leads to significant improvements in children's math scores by 0.0002 SD. The coefficient for literacy is statistically insignificant and quite small in magnitude. Table 12 also reports that, on average, 53.57 household toilets are constructed for every 1000 population in the district by 2018, with a stan-

dard deviation of 36.6, respectively. If we scale the effect by the average, our model predicts that constructing 53.57 toilets in the district through the Clean India Mission improves math scores by 0.109 SD with a standard error of 0.021, respectively.[11] The literacy score estimates are 0.015 SD with an error of 0.019. These results are plotted in Figure 10. The positive and significant math score improvements are qualitatively similar to the aggregated effects from Callaway and Sant'Anna (2021), and Wooldridge (2021). For the literacy scores, while we find positive changes in all the specifications, the point estimates are not consistently statistically significant.

6 Conclusion

We examine the effect of a large-scale sanitation program on children's cognitive abilities in India, using the staggered difference-in-differences approach from Callaway and Sant'Anna (2021) and find positive math and literacy score changes for children in the earliest treatment districts. We support the findings by estimating the group-specific effects, event-time effects, and the calendar-time effects of the program. We test for parallel trends using an aggregated event study design. Parallel trends violation is addressed using the honest difference-in-differences approach from Rambachan and Roth (2023).

We test for changes in children's school enrollment, the probability of households living in a thatched dwelling, and households' access to a toilet. These are the potential mechanisms through which the program can affect children's cognition. Additional robustness checks include finding similar improvements in raw, binary, and age-standardized math and literacy scores due to the program, and also estimating null associations between several district level characteristics and the program's implementation. This test provides support for non-random program placement. We also adopt the staggered difference-in-differences approach from Wooldridge (2021) and estimate dynamic program effects on children's cognition. Lastly, we examine the effect of the number of toilets constructed through the program and find positive changes in children's math and literacy

[11] These effects are scaled by multiplying the average values with the point estimates.

scores.

By examining the cognition benefits of the Clean India Mission, we contribute to a growing literature on the role of sanitation programs in improving children's education in developing countries. We also add to the limited research studying the effects of the program by demonstrating the positive benefits to children's education. The limitations of this work are that, due to lack of data we are unable to include information on the caste and religious groups of children, which would allow for examining these group-specific effects of the program. We are also unable to match the ASER and Clean India Mission datasets at a more granular level (e.g., block or panchayat) than the district. This could demonstrate a better picture of the program's effects. However, our findings credibly show that a program that is primarily targeted towards reducing open defecation and improving toilet-access and household living conditions, has generated improvements in children's cognition. This could potentially result in better human capital outcomes for these children in the future.

References

Abadie, A. (2005). Semiparametric difference-in-differences estimators. *The review of economic studies*, 72(1):1–19.

Adukia, A. (2017). Sanitation and education. *American Economic Journal: Applied Economics*, 9(2):23–59.

Adukia, A., Asher, S., and Novosad, P. (2020). Educational investment responses to economic opportunity: evidence from indian road construction. *American Economic Journal: Applied Economics*, 12(1):348–376.

Aizer, A., Currie, J., Simon, P., and Vivier, P. (2018). Do low levels of blood lead reduce children's future test scores? *American Economic Journal: Applied Economics*, 10(1):307–341.

Alexander, K., Allton, C., Felsman, C., Hahn, M., Okegbe, T., Palmer, D., Taglieri, J., Thibert, M., et al. (2016). Ending open defecation in india: Insights on implementation and behavior change for swachh bharat abhiyan.

Alsan, M. and Goldin, C. (2019). Watersheds in child mortality: The role of effective water and sewerage infrastructure, 1880–1920. *Journal of Political Economy*, 127(2):586–638.

Andres, L. A., Deb, S., Joseph, G., Larenas, M. I., and Grabinsky Zabludovsky, J. (2020). A multiple-arm, cluster-randomized impact evaluation of the clean india (swachh bharat) mission program in rural punjab, india. *World Bank Policy Research Working Paper*, (9249).

Aquino, P. G., Justin, M. S. M., and Jalagat, R. C. (2021). Clean india mission and its impact on cities of tourist importance in india. *Strategies in Sustainable Tourism, Economic Growth and Clean Energy*, pages 89–107.

Arnold, B. F., Null, C., Luby, S. P., Unicomb, L., Stewart, C. P., Dewey, K. G., Ahmed, T., Ashraf, S., Christensen, G., Clasen, T., et al. (2013). Cluster-randomised controlled trials of individual and combined water, sanitation, hygiene and nutritional interventions in rural bangladesh and kenya: the wash benefits study design and rationale. *BMJ open*, 3(8):e003476.

Athey, S., Imbens, G. W., Wager, S., et al. (2016). Efficient inference of average treatment effects in high dimensions via approximate residual balancing. Technical report.

Augsburg, B., Baquero, J. P., Gautam, S., and Rodriguez-Lesmes, P. (2023). Sanitation and marriage markets in india: Evidence from the total sanitation campaign. *Journal of Development Economics*, 163:103092.

Augsburg, B. and Rodriguez-Lesmes, P. A. (2018). Sanitation and child health in india. *World Development*, 107:22–39.

Azurin, J. and Alvero, M. (1974). Field evaluation of environmental sanitation measures against cholera. *Bulletin of the World Health Organization*, 51(1):19.

Banerjee, T., Kumar, M., Mall, R., and Singh, R. (2017). Airing 'clean air'in clean india mission. *Environmental Science and Pollution Research*, 24(7):6399–6413.

Barham, T. (2012). Enhancing cognitive functioning: Medium-term effects of a health and family planning program in matlab. *American Economic Journal: Applied Economics*, 4(1):245–273.

Barrett, P., Treves, A., Shmis, T., and Ambasz, D. (2019). The impact of school infrastructure on learning: A synthesis of the evidence.

Battiston, D., Cruces, G., Lopez-Calva, L. F., Lugo, M. A., and Santos, M. E. (2013). Income and beyond: Multidimensional poverty in six latin american countries. *Social indicators research*, 112:291–314.

Behera, M. R., Pradhan, H. S., Behera, D., Jena, D., and Satpathy, S. K. (2021). Achievements and challenges of india's sanitation campaign under clean india mission: A commentary. *Journal of Education and Health Promotion*, 10.

Bekele, T., Rawstorne, P., and Rahman, B. (2020). Effect of water, sanitation and hygiene interventions alone and combined with nutrition on child growth in low and middle income countries: a systematic review and meta-analysis. *BMJ open*, 10(7):e034812.

Belmonte, A., Bove, V., D'Inverno, G., and Modica, M. (2017). School infrastructure spending and educational outcomes in northern italy.

Berkman, D. S., Lescano, A. G., Gilman, R. H., Lopez, S. L., and Black, M. M. (2002). Effects of stunting, diarrhoeal disease, and parasitic infection during infancy on cognition in late childhood: a follow-up study. *The Lancet*, 359(9306):564–571.

Berner, M. M. (1993). Building conditions, parental involvement, and student achievement in the district of columbia public school system. *Urban Education*, 28(1):6–29.

Bharadwaj, P., Gibson, M., Zivin, J. G., and Neilson, C. (2017). Gray matters: Fetal pollution exposure and human capital formation. *Journal of the Association of Environmental and Resource Economists*, 4(2):505–542.

Bhat, D. (2017). "harbinger of a new era"? evaluating the effect of india's right to education act on learning outcomes.'. *M-RCBH Associate Working Paper*, 76.

Bhattacharya, N., Pakrashi, D., Saha, S., and Sahoo, S. (2023). Impact of early childhood school intervention on enrollment and learning outcomes: Evaluation of a public program in india.

Bhattacharya, S., Sharma, D., and Sharma, P. (2018). Swachh bharat mission: an integrative approach to attain public health in india. *International Journal of Environment and Health*, 9(2):197–212.

Birdthistle, I., Dickson, K., Freeman, M., and Javidi, L. (2011). What impact does the provision of separate toilets for girls at schools have on their primary and secondary school enrolment, attendance and completion? a systematic review of the evidence. *Social Science Research Unit, Institute of Education, University of London*, 6.

Borusyak, K., Jaravel, X., and Spiess, J. (2021). Revisiting event study designs: Robust and efficient estimation. *arXiv preprint arXiv:2108.12419*.

Branham, D. (2004). The wise man builds his house upon the rock: The effects of inadequate school building infrastructure on student attendance. *Social science quarterly*, 85(5):1112–1128.

Buttenheim, A. M. (2008). The sanitation environment in urban slums: implications for child health. *Population and environment*, 30:26–47.

Callaway, B. and Sant'Anna, P. H. (2021). Difference-in-differences with multiple time periods. *Journal of econometrics*, 225(2):200–230.

Cameron, L., Chase, C., Haque, S., Joseph, G., Pinto, R., and Wang, Q. (2021). Childhood stunting and cognitive effects of water and sanitation in indonesia. *Economics & Human Biology*, 40:100944.

Cameron, L. A., Shah, M., and Olivia, S. (2013). Impact evaluation of a large-scale rural sanitation project in indonesia. *World Bank policy research working paper*, (6360).

Chakrabarti, S., Gune, S., Bruckner, T. A., Strominger, J., and Singh, P. (2024). Toilet construction under the swachh bharat mission and infant mortality in india. *Scientific Reports*, 14(1):20340.

Chakraborty, T. and Jayaraman, R. (2019). School feeding and learning achievement: Evidence from india's midday meal program. *Journal of Development Economics*, 139:249–265.

Chaudhury, N., Christiaensen, L., and Asadullah, M. N. (2006). Schools, household, risk, and gender: Determinants of child schooling in ethiopia.

Clarke, D. and Tapia-Schythe, K. (2021). Implementing the panel event study. *The Stata Journal*, 21(4):853–884.

Coffey, D., Geruso, M., and Spears, D. (2018). Sanitation, disease externalities and anaemia: evidence from nepal. *The Economic Journal*, 128(611):1395–1432.

Cronk, R., Slaymaker, T., and Bartram, J. (2015). Monitoring drinking water, sanitation, and hygiene in non-household settings: priorities for policy and practice. *International Journal of Hygiene and Environmental Health*, 218(8):694–703.

Cuesta, A., Glewwe, P., and Krause, B. (2016). School infrastructure and educational outcomes: a literature review, with special reference to latin america. *Economia*, 17(1):95–130.

Cuesta, J. (2007). Child malnutrition and the provision of water and sanitation in the philippines. *Journal of the Asia Pacific Economy*, 12(2):125–157.

Cunningham, S. (2021). *Causal inference: The mixtape*. Yale university press.

Curtis, V. (2019). Explaining the outcomes of the'clean india'campaign: institutional behaviour and sanitation transformation in india. *BMJ global health*, 4(5):e001892.

Cutler, D., Deaton, A., and Lleras-Muney, A. (2006). The determinants of mortality. *Journal of economic perspectives*, 20(3):97–120.

Dandabathula, G., Bhardwaj, P., Burra, M., Rao, P. V. P., and Rao, S. S. (2019). Impact assessment of india's swachh bharat mission–clean india campaign on acute diarrheal disease outbreaks: Yes, there is a positive change. *Journal of family medicine and primary care*, 8(3):1202.

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

de Barros, A., Fajardo-Gonzalez, J., Glewwe, P., and Sankar, A. (2022). Large-scale efforts to improve teaching and child learning: Experimental evidence from india. *Journal of Development Economics (conditionally accepted via pre-results review)*.

De Chaisemartin, C. and d'Haultfoeuille, X. (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, 110(9):2964–2996.

Dearden, K. A., Brennan, A. T., Behrman, J. R., Schott, W., Crookston, B. T., Humphries, D. L., Penny, M. E., and Fernald, L. C. (2017). Does household access to improved water and sanitation in infancy and childhood predict better vocabulary test performance in ethiopian, indian, peruvian and vietnamese cohort studies? *BMJ open*, 7(3):e013201.

Dibaba, D. B., Alemu, B. M., and Debela, S. A. (2024). Impact of water, sanitation, and hygiene (wash) interventions on gender-specific school attendance and learning outcomes: A systematic review and meta-analysis protocol. *Plos one*, 19(8):e0308144.

Dickinson, K. L., Patil, S. R., Pattanayak, S. K., Poulos, C., and Yang, J.-H. (2015). Nature's call: impacts of sanitation choices in orissa, india. *Economic Development and Cultural Change*, 64(1):1–29.

Dix-Cooper, L., Eskenazi, B., Romero, C., Balmes, J., and Smith, K. R. (2012). Neurodevelopmental performance among school age children in rural guatemala is associated with prenatal and postnatal exposure to carbon monoxide, a marker for exposure to woodsmoke. *Neurotoxicology*, 33(2):246–254.

Dreibelbis, R., Freeman, M. C., Greene, L. E., Saboori, S., and Rheingans, R. (2014). The impact of school water, sanitation, and hygiene interventions on the health of younger siblings of pupils: a cluster-randomized trial in kenya. *American journal of public health*, 104(1):e91–e97.

Dreibelbis, R., Greene, L. E., Freeman, M. C., Saboori, S., Chase, R. P., and Rheingans, R. (2013a). Water, sanitation, and primary school attendance: A multi-level assessment of determinants of household-reported absence in kenya. *International Journal of Educational Development*, 33(5):457–465.

Dreibelbis, R., Winch, P. J., Leontsini, E., Hulland, K. R., Ram, P. K., Unicomb, L., and Luby, S. P. (2013b). The integrated behavioural model for water, sanitation, and hygiene: a systematic review of behavioural models and a framework for designing and evaluating behaviour change interventions in infrastructure-restricted settings. *BMC public health*, 13(1):1–13.

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

Durán-Narucki, V. (2008). School building condition, school attendance, and academic achievement in new york city public schools: A mediation model. *Journal of environmental psychology*, 28(3):278–286.

Elledge, M. F. and McClatchey, M. (2013). *India, urban sanitation, and the toilet challenge*. RTI International/RTI Press.

Fentiman, A., Hall, A., and Bundy, D. (1999). School enrolment patterns in rural ghana: a comparative study of the impact of location, gender, age and health on children's access to basic schooling. *Comparative education*, 35(3):331–349.

Fink, G., Günther, I., and Hill, K. (2011). The effect of water and sanitation on child health: evidence from the demographic and health surveys 1986–2007. *International journal of epidemiology*, 40(5):1196–1204.

Fischer Walker, C. L., Lamberti, L., Adair, L., Guerrant, R. L., Lescano, A. G., Martorell, R., Pinkerton, R. C., and Black, R. E. (2012). Does childhood diarrhea influence cognition beyond the diarrhea-stunting pathway? *PloS one*, 7(10):e47908.

Flutter, J. (2006). 'this place could help you learn': student participation in creating better school environments. *Educational review*, 58(2):183–193.

Fotio, H. K. and Nguea, S. M. (2022). Access to water and sanitation in africa: Does globalization matter? *International Economics*, 170:79–91.

Freeman, M., Clasen, T., Dreibelbis, R., Saboori, S., Greene, L., Brumback, B., Muga, R., and Rheingans, R. (2014). The impact of a school-based water supply and treatment, hygiene, and sanitation programme on pupil diarrhoea: a cluster-randomized trial. *Epidemiology & Infection*, 142(2):340–351.

Freeman, M. C., Greene, L. E., Dreibelbis, R., Saboori, S., Muga, R., Brumback, B., and Rheingans, R. (2012). Assessing the impact of a school-based water treatment, hygiene and sanitation programme on pupil absence in nyanza province, kenya: a cluster-randomized trial. *Tropical medicine & international health*, 17(3):380–391.

Freyaldenhoven, S., Hansen, C., and Shapiro, J. M. (2019). Pre-event trends in the panel event-study design. *American Economic Review*, 109(9):3307–38.

Gevrek, Z. E., Neumeier, C., and Gevrek, D. (2018). Explaining the gender test score gap in mathematics: The role of gender inequality. *Available at SSRN 3111114*.

Ghosh, A. and Cairncross, S. (2014). The uneven progress of sanitation in india. *Journal of Water, Sanitation and Hygiene for Development*, 4(1):15–22.

Gilmartin, A. A. and Petri Jr, W. A. (2015). Exploring the role of environmental enteropathy in malnutrition, infant development and oral vaccine response. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 370(1671):20140143.

Glick, P. (2008). What policies will reduce gender schooling gaps in developing countries: Evidence and interpretation. *World Development*, 36(9):1623–1646.

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, 225(2):254–277.

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

Grantham-McGregor, S. M., Pollitt, E., Wachs, T. D., Meisels, S. J., and Scott, K. G. (1999). Summary of the scientific evidence on the nature and determinants of child development and their implications for programmatic interventions with young children. *Food and nutrition bulletin*, 20(1):4–6.

Greenberg, M. T., Lengua, L. J., Coie, J. D., Pinderhughes, E. E., Bierman, K., Dodge, K. A., Lochman, J. E., and McMahon, R. J. (1999). Predicting developmental outcomes at school entry using a multiple-risk model: four american communities. *Developmental psychology*, 35(2):403.

Hathaway, W. E. (1995). Effects of school lighting on physical development and school performance. *The Journal of Educational Research*, 88(4):228–242.

Headey, D. and Palloni, G. (2019). Water, sanitation, and child health: evidence from subnational panel data in 59 countries. *Demography*, 56(2):729–752.

Heckman, J. J., Ichimura, H., and Todd, P. (1998). Matching as an econometric evaluation estimator. *The review of economic studies*, 65(2):261–294.

Heckman, J. J., Ichimura, H., and Todd, P. E. (1997). Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme. *The review of economic studies*, 64(4):605–654.

Hermann, Z. and Kopasz, M. (2021). Educational policies and the gender gap in test scores: A cross-country analysis. *Research Papers in Education*, 36(4):461–482.

Hossain, M. A., Mahajan, K., and Sekhri, S. (2022). Access to toilets and violence against women. *Journal of Environmental Economics and Management*, 114:102695.

Hueso, A. and Bell, B. (2013). An untold story of policy failure: the total sanitation campaign in india. *Water Policy*, 15(6):1001–1017.

Hutton, G. (2013). Global costs and benefits of reaching universal coverage of sanitation and drinking-water supply. *Journal of water and health*, 11(1):1–12.

Hutton, G., Patil, S., Kumar, A., Osbert, N., and Odhiambo, F. (2020). Comparison of the costs and benefits of the clean india mission. *World Development*, 134:105052.

Jalan, J. and Ravallion, M. (2003). Does piped water reduce diarrhea for children in rural india? *Journal of econometrics*, 112(1):153–173.

Jayachandran, S. (2015). The roots of gender inequality in developing countries. *economics*, 7(1):63–88.

Jha, N. (2010). Access of the poor to water supply and sanitation in india: salient concepts, issues and cases. Technical report, working paper.

Karmakar, A. and Villa, K. M. (2022). Going to school at better schools can improve children's health: The health benefits of the 2009 right to education act in india.

Khandelwal, S., Tamboli, P., Madhup, S., and Dandabathula, G. (2020). Assessment of swachh bharat mission-clean india campaign's contribution in combating violence against women. *International Journal of Gender and Women's Studies*, 8(2):143–154.

Kumar, A. (2017). Beyond toilets and targets: sanitation mission in india. *Development in Practice*, 27(3):408–413.

Kumar, S. and Vollmer, S. (2013). Does access to improved sanitation reduce childhood diarrhea in rural india? *Health Economics*, 22(4):410–427.

Lavy, V. (2015). Do differences in schools' instruction time explain international achievement gaps? evidence from developed and developing countries. *The Economic Journal*, 125(588):F397–F424.

Liu, J. and Lewis, G. (2014). Environmental toxicity and poor cognitive outcomes in children and adults. *Journal of environmental health*, 76(6):130.

Lokshin, M. and Yemtsov, R. (2005). Has rural infrastructure rehabilitation in georgia helped the poor? *The World Bank Economic Review*, 19(2):311–333.

Luby, S. (2014). Is targeting access to sanitation enough? *The Lancet Global Health*, 2(11):e619–e620.

Machin, S. and Pekkarinen, T. (2008). Global sex differences in test score variability. *Science*, 322(5906):1331–1332.

Mansuri, G., Sami, M. F., Ali, M., Doan, H. T. T., Javed, B., Pandey, P., and Asia, S. (2018). When water becomes a hazard: A diagnostic report on the state of water supply, sanitation and poverty in pakistan and its impact on child stunting. *WASH Poverty Diagnostic Series. Washington, DC: World Bank Group*.

Mara, D., Lane, J., Scott, B., and Trouba, D. (2010). Sanitation and health. *PLoS medicine*, 7(11):e1000363.

Mathew, K., Zachariah, S., Shordt, K., Snel, M., Cairncross, S., Biran, A., and Schmidt, W.-P. (2009). The sustainability and impact of school sanitation, water and hygiene education in southern india. *Waterlines*, pages 275–292.

McGowen, R. S. (2007). *The impact of school facilities on student achievement, attendance, behavior, completion rate and teacher turnover rate in selected Texas high schools*. Texas A&M University.

Miller, S., Johnson, N., and Wherry, L. R. (2021). Medicaid and mortality: new evidence from linked survey and administrative data. *The Quarterly Journal of Economics*, 136(3):1783–1829.

Mohapatra, G. (2019). Projected behavioural change in swachh bharat mission: a public policy perspective. *Indian Journal of Public Administration*, 65(2):451–474.

Momberg, D. J., Voth-Gaeddert, L. E., Richter, L. M., Norris, S. A., and Said-Mohamed, R. (2022). Rethinking water, sanitation, and hygiene for human growth and development. *Global Public Health*, pages 1–10.

Muralidharan, K. and Sheth, K. (2016). Bridging education gender gaps in developing countries: The role of female teachers. *Journal of Human Resources*, 51(2):269–297.

Nagengast, A., Rios-Avila, F., and Yotov, Y. (2024). The european single market and intra-eu trade: An assessment with heterogeneity-robust difference-in-differences methods. Technical report, LeBow College of Business, Drexel University.

Nandi, A., Megiddo, I., Ashok, A., Verma, A., and Laxminarayan, R. (2017). Reduced burden of childhood diarrheal diseases through increased access to water and sanitation in india: A modeling analysis. *Social Science & Medicine*, 180:181–192.

Njuguna, J. (2016). Effect of eliminating open defecation on diarrhoeal morbidity: an ecological study of nyando and nambale sub-counties, kenya. *BMC Public Health*, 16(1):1–6.

Null, C., Stewart, C. P., Pickering, A. J., Dentz, H. N., Arnold, B. F., Arnold, C. D., Benjamin-Chung, J., Clasen, T., Dewey, K. G., Fernald, L. C., et al. (2018). Effects of water quality, sanitation, handwashing, and nutritional interventions on diarrhoea and child growth in rural kenya: a cluster-randomised controlled trial. *The Lancet Global Health*, 6(3):e316–e329.

Orgill-Meyer, J. and Pattanayak, S. K. (2020). Improved sanitation increases long-term cognitive test scores. *World Development*, 132:104975.

Ortiz-Correa, J. S., Resende Filho, M., and Dinar, A. (2016). Impact of access to water and sanitation services on educational attainment. *Water Resources and Economics*, 14:31–43.

Oster, E. and Thornton, R. (2011). Menstruation, sanitary products, and school attendance: evidence from a randomized evaluation. *American Economic Journal: Applied Economics*, 3(1):91–100.

Pakhtigian, E. L., Dickinson, K. L., Orgill-Meyer, J., and Pattanayak, S. K. (2022). Sustaining latrine use: Peers, policies, and sanitation behaviors. *Journal of Economic Behavior & Organization*, 200:223–242.

Parnwell, R. N. (2015). *Influence of school infrastructure on academic performance in Public primary schools in Ruiri location-Meru county, Kenya*. PhD thesis, University of Nairobi.

Patil, S. R., Arnold, B. F., Salvatore, A. L., Briceno, B., Ganguly, S., Colford Jr, J. M., and Gertler, P. J. (2014). The effect of india's total sanitation campaign on defecation behaviors and child health in rural madhya pradesh: a cluster randomized controlled trial. *PLoS medicine*, 11(8):e1001709.

Patrinos, H. A. (2000). Market forces in education. *European Journal of Education*, 35(1):61–80.

Paxson, C. and Schady, N. R. (2002). The allocation and impact of social funds: spending on school infrastructure in peru. *The World Bank Economic Review*, 16(2):297–319.

Petri, W. A., Miller, M., Binder, H. J., Levine, M. M., Dillingham, R., Guerrant, R. L., et al. (2008). Enteric infections, diarrhea, and their impact on function and development. *The Journal of clinical investigation*, 118(4):1277–1290.

Pickering, A. J., Djebbari, H., Lopez, C., Coulibaly, M., and Alzua, M. L. (2015). Effect of a community-led sanitation intervention on child diarrhoea and child growth in rural mali: a cluster-randomised controlled trial. *The Lancet Global Health*, 3(11):e701–e711.

Prithviraj, R., Sivaprakasam, P., Mahendran, C., Norman, P., Pandyan, K., et al. (2017). Status of sanitation in the rural areas of a health unit dis-trict, tamil nadu, india. *National Journal of Community Medicine*, 8(09):521–525.

Rahman, M. H. U., Malik, M. A., Chauhan, S., Patel, R., Singh, A., and Mittal, A. (2020). Examining the linkage between open defecation and child malnutrition in india. *Children and Youth Services Review*, 117:105345.

Rambachan, A. and Roth, J. (2023). A more credible approach to parallel trends. *Review of Economic Studies*, 90(5):2555–2591.

Rauniyar, G., Orbeta Jr, A., and Sugiyarto, G. (2011). Impact of water supply and sanitation assistance on human welfare in rural pakistan. *Journal of development effectiveness*, 3(1):62–102.

Rosado, J. L., Ronquillo, D., Kordas, K., Rojas, O., Alatorre, J., Lopez, P., Garcia-Vargas, G., del Carmen Caamaño, M., Cebrián, M. E., and Stoltzfus, R. J. (2007). Arsenic exposure and cognitive performance in mexican schoolchildren. *Environmental health perspectives*, 115(9):1371–1375.

Roth, J. (2022). Pretest with caution: Event-study estimates after testing for parallel trends. *American Economic Review: Insights*, 4(3):305–22.

Routray, P., Torondel, B., Jenkins, M. W., Clasen, T., and Schmidt, W.-P. (2017). Processes and challenges of community mobilisation for latrine promotion under nirmal bharat abhiyan in rural odisha, india. *BMC Public Health*, 17:1–15.

Ruff, H. A., Markowitz, M. E., Bijur, P. E., and Rosen, J. F. (1996). Relationships among blood lead levels, iron deficiency, and cognitive development in two-year-old children. *Environmental Health Perspectives*, 104(2):180–185.

Saenz, J. L., Wong, R., and Ailshire, J. A. (2018). Indoor air pollution and cognitive function among older mexican adults. *J Epidemiol Community Health*, 72(1):21–26.

Santos, D. N., Assis, A. M. O., Bastos, A. C. S., Santos, L. M., Santos, C. A. S., Strina, A., Prado, M. S., Almeida-Filho, N. M., Rodrigues, L. C., and Barreto, M. L. (2008). Determinants of cognitive function in childhood: a cohort study in a middle income context. *BMC Public health*, 8:1–15.

Sant'Anna, P. H. and Zhao, J. (2020). Doubly robust difference-in-differences estimators. *Journal of econometrics*, 219(1):101–122.

Schady, N. and Paxson, C. H. (1999). Do school facilities matter? the case of the peruvian social fund (foncodes). *The Case of the Peruvian Social Fund (Foncodes)(November 1999)*.

Schmidt, W.-P. (2014). The elusive effect of water and sanitation on the global burden of disease.

Sclar, G., Garn, J., Penakalapati, G., Alexander, K., Krauss, J., Freeman, M., Boisson, S., Medlicott, K., and Clasen, T. (2017). Effects of sanitation on cognitive development and school absence: a systematic review. *International Journal of Hygiene and Environmental Health*, 220(6):917–927.

Scriptore Moreira, J. S., Azzoni, C. R., and Menezes-Filho, N. (2024). Sanitation conditions and education outcomes in brazilian municipalities. *Sustainable Development*, 32(3):1987–2000.

Shah, M. and Steinberg, B. (2019). The right to education act: Trends in enrollment, test scores, and school quality. In *AEA Papers and Proceedings*, volume 109, pages 232–38.

Sharma, M. K., Adhikari, R., van Tejlingen, E., Devkota, B., and Khanal, S. P. (2024). Improved water, sanitation and hygiene facilities at school and their effect on educational achievement in basic level students in nepal. *International Journal of Health Promotion and Education*, pages 1–12.

Spears, D. (2012). Effects of rural sanitation on infant mortality and human capital: Evidence from india's total sanitation campaign. *Washington, DC: Princeton University*.

Spears, D. (2013). How much international variation in child height can sanitation explain? *World Bank policy research working paper*, (6351).

Spears, D. and Lamba, S. (2016). Effects of early-life exposure to sanitation on childhood cognitive skills: Evidence from india's total sanitation campaign. *Journal of Human Resources*, 51(2):298–327.

Stewart, C. P., Kariger, P., Fernald, L., Pickering, A. J., Arnold, C. D., Arnold, B. F., Hubbard, A. E., Dentz, H. N., Lin, A., Meerkerk, T. J., et al. (2018). Effects of water quality, sanitation, handwashing, and nutritional interventions on child development in rural kenya (wash benefits kenya): a cluster-randomised controlled trial. *The lancet child & adolescent health*, 2(4):269–280.

Strauss, J. and Thomas, D. (1998). Health, nutrition, and economic development. *Journal of economic literature*, 36(2):766–817.

Sun, L. and Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, 225(2):175–199.

Thakur, R., Singh, B. B., Jindal, P., Aulakh, R. S., and Gill, J. P. S. (2018). The clean india mission: Public and animal health benefits. *Acta tropica*, 186:5–15.

Tofail, F., Fernald, L. C., Das, K. K., Rahman, M., Ahmed, T., Jannat, K. K., Unicomb, L., Arnold, B. F., Ashraf, S., Winch, P. J., et al. (2018). Effect of water quality, sanitation, hand washing, and nutritional interventions on child development in rural bangladesh (wash benefits bangladesh): a cluster-randomised controlled trial. *The Lancet Child & Adolescent Health*, 2(4):255–268.

UNICEF et al. (2018). Financial and economic impacts of the swachh bharat mission in india. *UNICEF: New Delhi, India*.

Vikram, K. and Chindarkar, N. (2020). Bridging the gaps in cognitive achievement in india: The crucial role of the integrated child development services in early childhood. *World Development*, 127:104697.

Villa, K. M. (2017). Multidimensional human capital formation in a developing country: Health, cognition and locus of control in the philippines. *Economics & Human Biology*, 27:184–197.

Vogl, T. S. (2014). Height, skills, and labor market outcomes in mexico. *Journal of Development Economics*, 107:84–96.

Wooldridge, J. M. (2021). Two-way fixed effects, the two-way mundlak regression, and difference-in-differences estimators. *Available at SSRN 3906345*.

Wooldridge, J. M. (2023). Simple approaches to nonlinear difference-in-differences with panel data. *The Econometrics Journal*, 26(3):C31–C66.

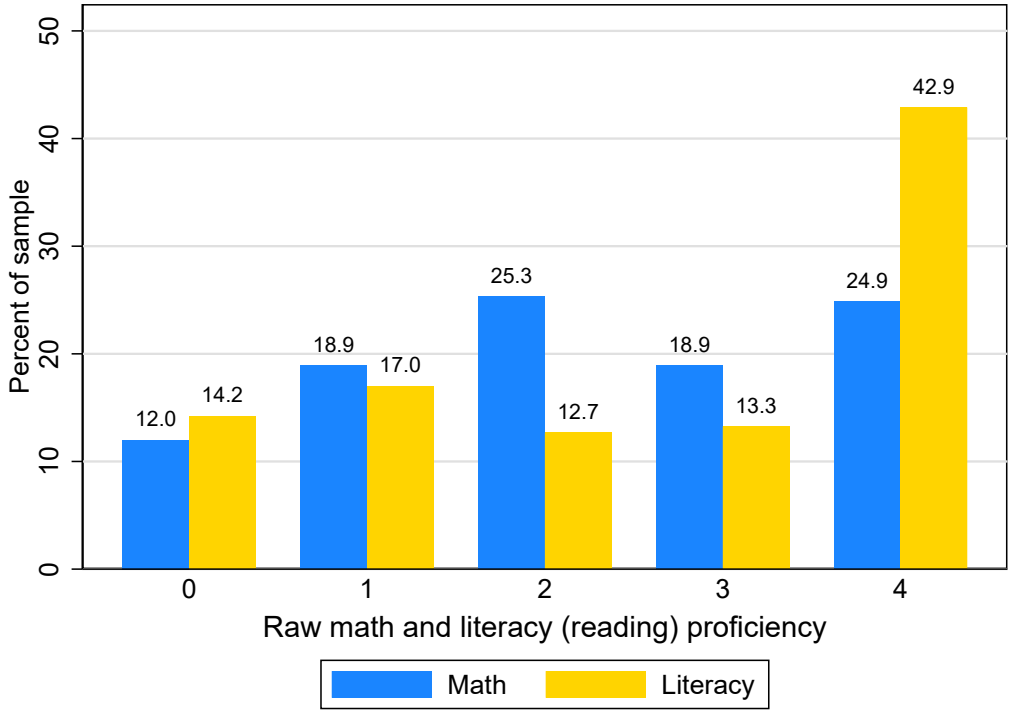


Figure 1: Percentage of children in each level of the math and literacy tests employed by ASER. The literacy assessment has 5 levels: cannot read (0), can read letters (1), can read words (2), read grade 1 text (3), and read grade 2 text (4). The five assessment levels for math include cannot do basic arithmetic (0), can recognize single-digit numbers (2), can recognize double-digit numbers (3), can perform two-digit subtraction (3), and can perform three-by-one-digit division (4).

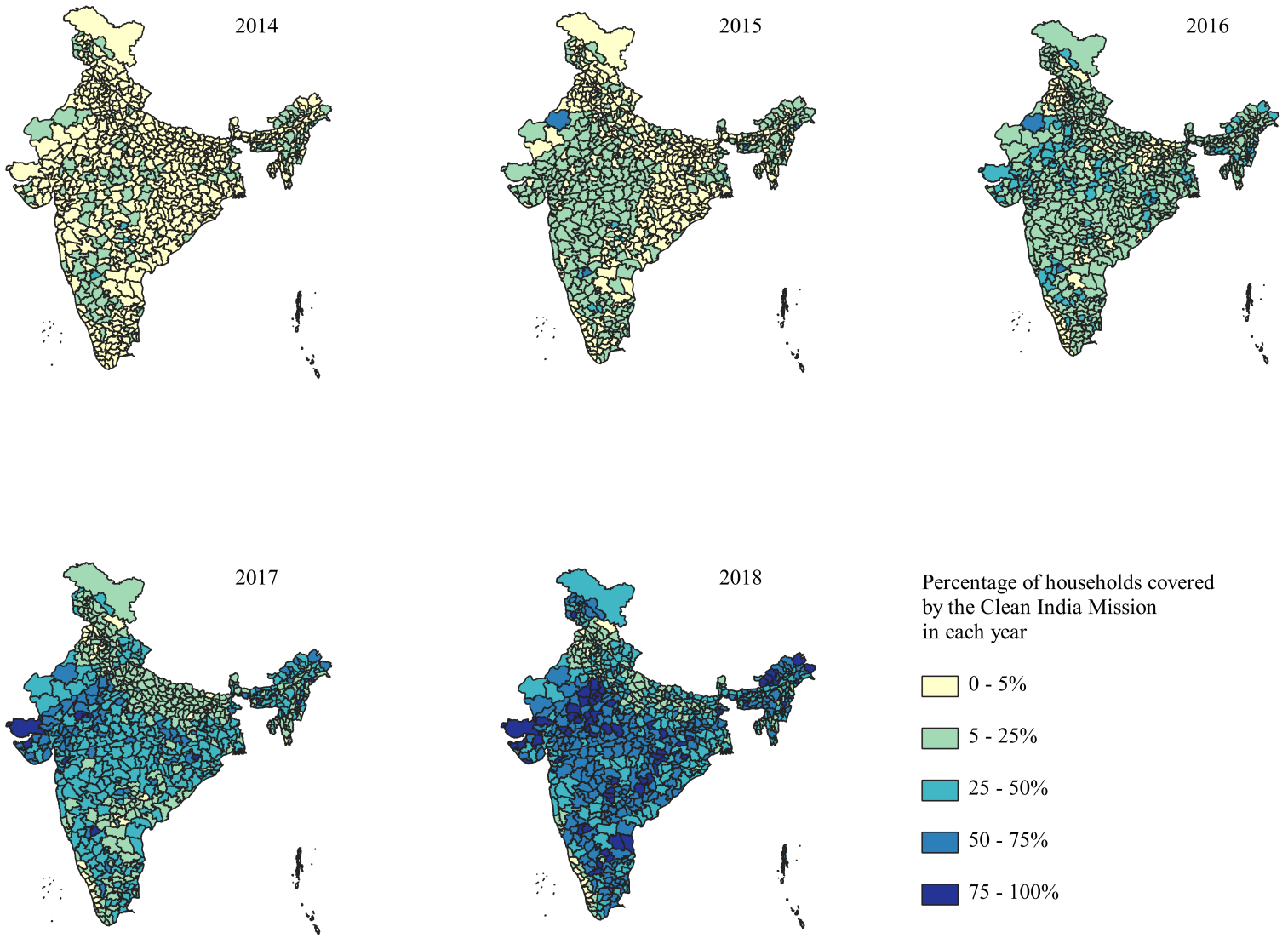


Figure 2: Percentage of households covered by the Clean India Mission in each district between 2014 and 2018. Source: District level data from the Government of India's Ministry of Drinking Water and Sanitation.

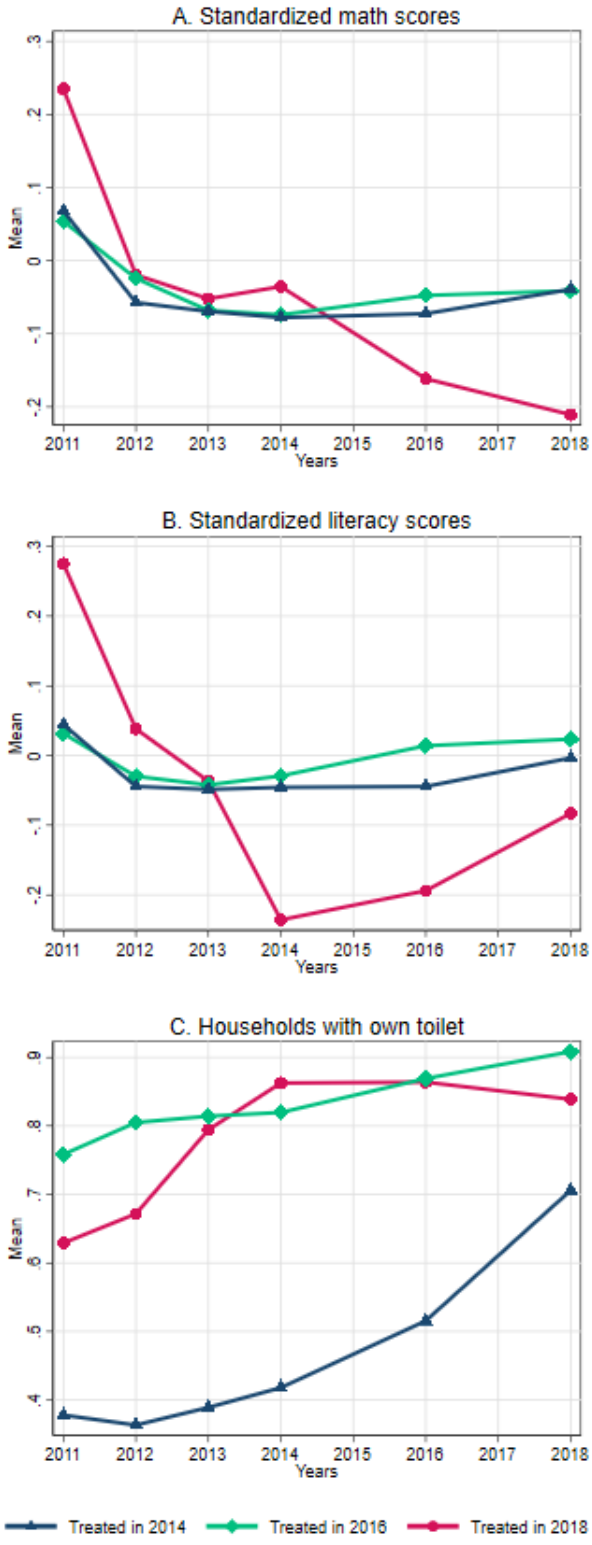


Figure 3: Trends in standardized math scores, standardized literacy scores, and percentage of households with a toilet. Notes: Standardized math scores shown in Panel A. Standardized math scores shown in Panel B. Scores are standardized using the district-specific mean and standard deviation over the entire duration of the ASER surveys (2011 - 2018). Panel C shows the percentage of households with a toilet.

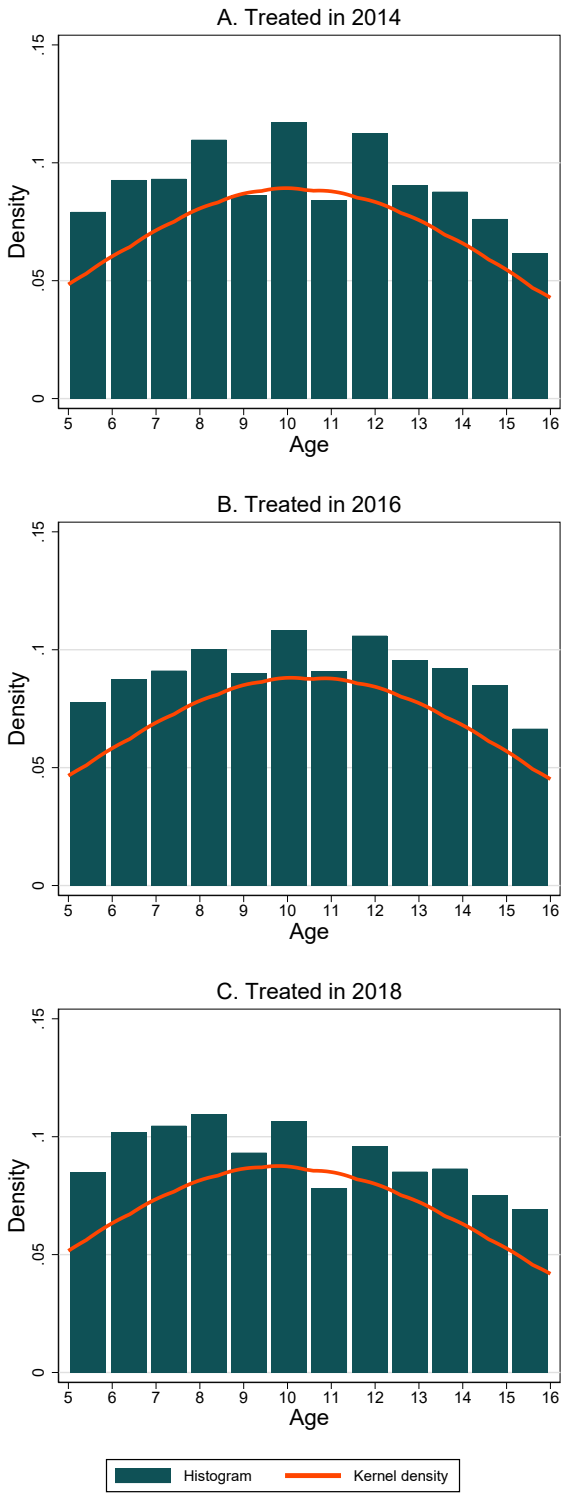


Figure 4: Distribution of age of sample children in districts first treated in 2014 (Panel A), first treated in 2016 (Panel B), and first treated in 2018 (Panel C).

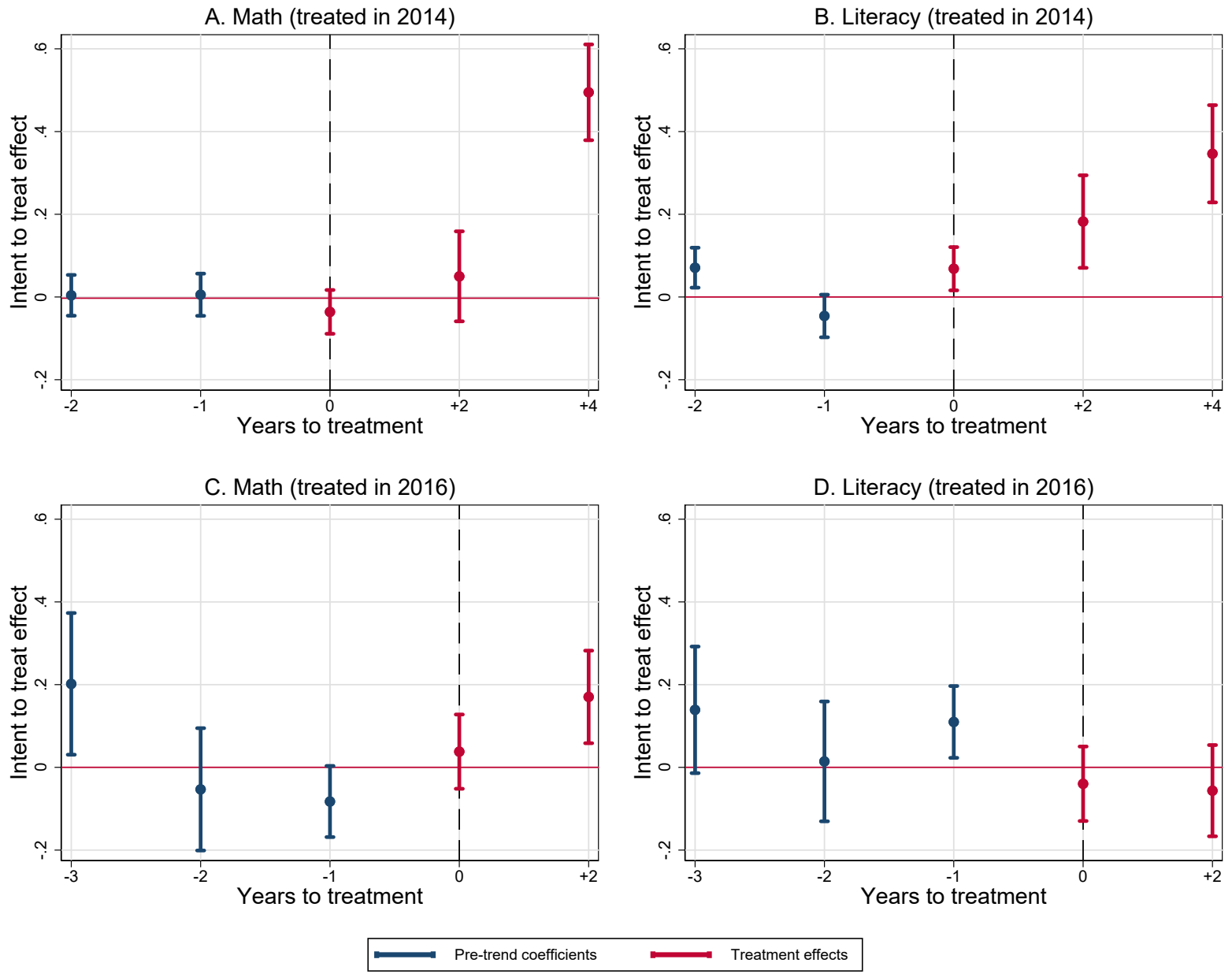


Figure 5: Math and literacy group-time intent to treat effects of Clean India Mission using Callaway and Sant'Anna (2021). Notes: Lag year coefficients are shown in red. Base year for each lag coefficient is the year just prior to first treatment. Lead coefficients are shown in blue, and the base year for each lead coefficient is the respective previous year. Parallel trends holds if all pre-treatment coefficients (lead coefficients) are statistically similar to zero. Each bar indicates the 95% confidence intervals and includes the point estimates.

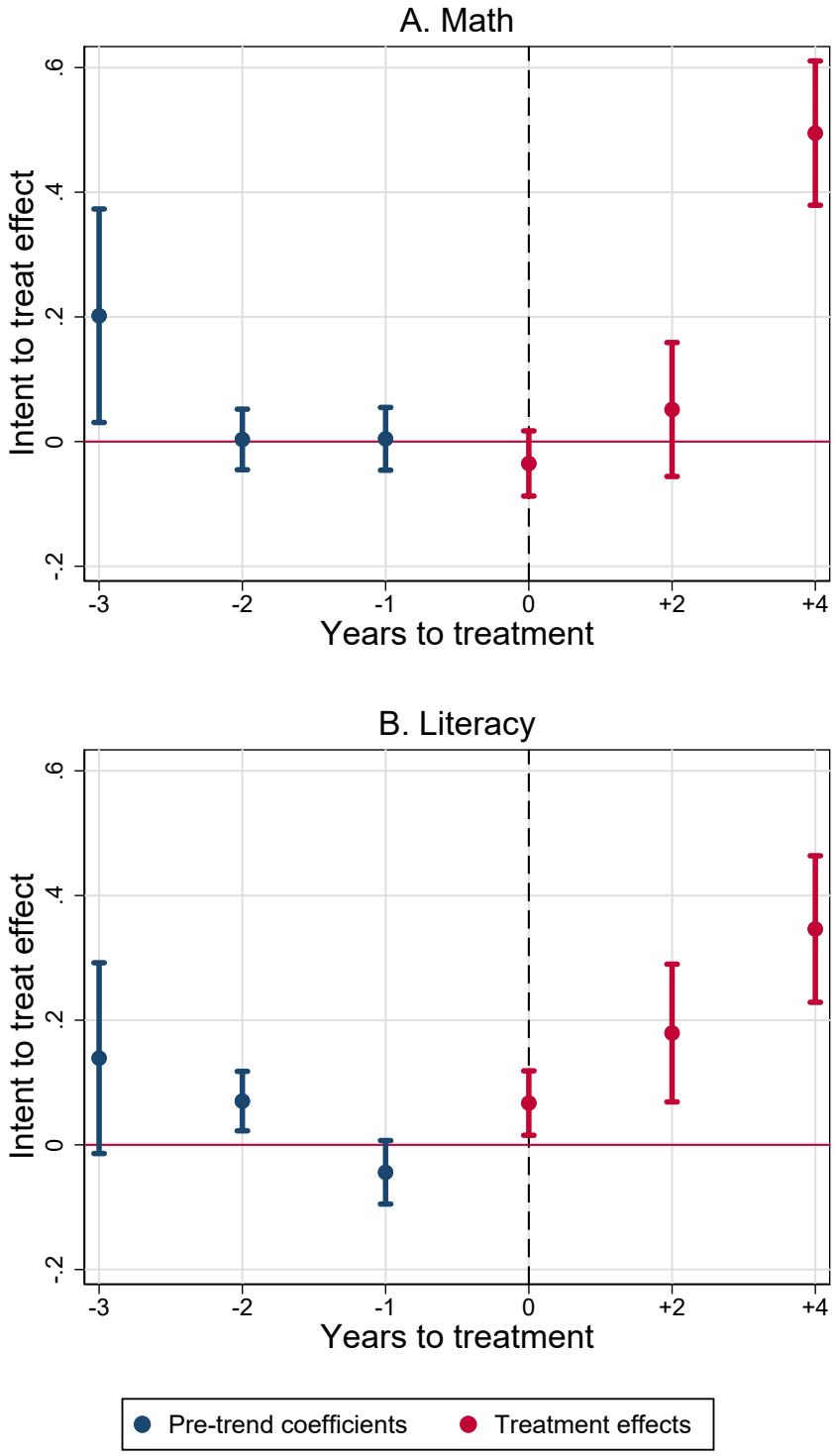


Figure 6: Event study using the doubly robust difference-in-differences approach by Callaway and Sant'Anna (2021). Notes: Lag year coefficients are shown in red. Base year for each lag coefficient is the year just prior to first treatment. Lead coefficients are shown in blue, and the base year for each lead coefficient is the respective previous year. Parallel trends holds if all pre-treatment coefficients (lead coefficients) are statistically similar to zero. Each bar indicates the 95% confidence intervals and includes the point estimates.

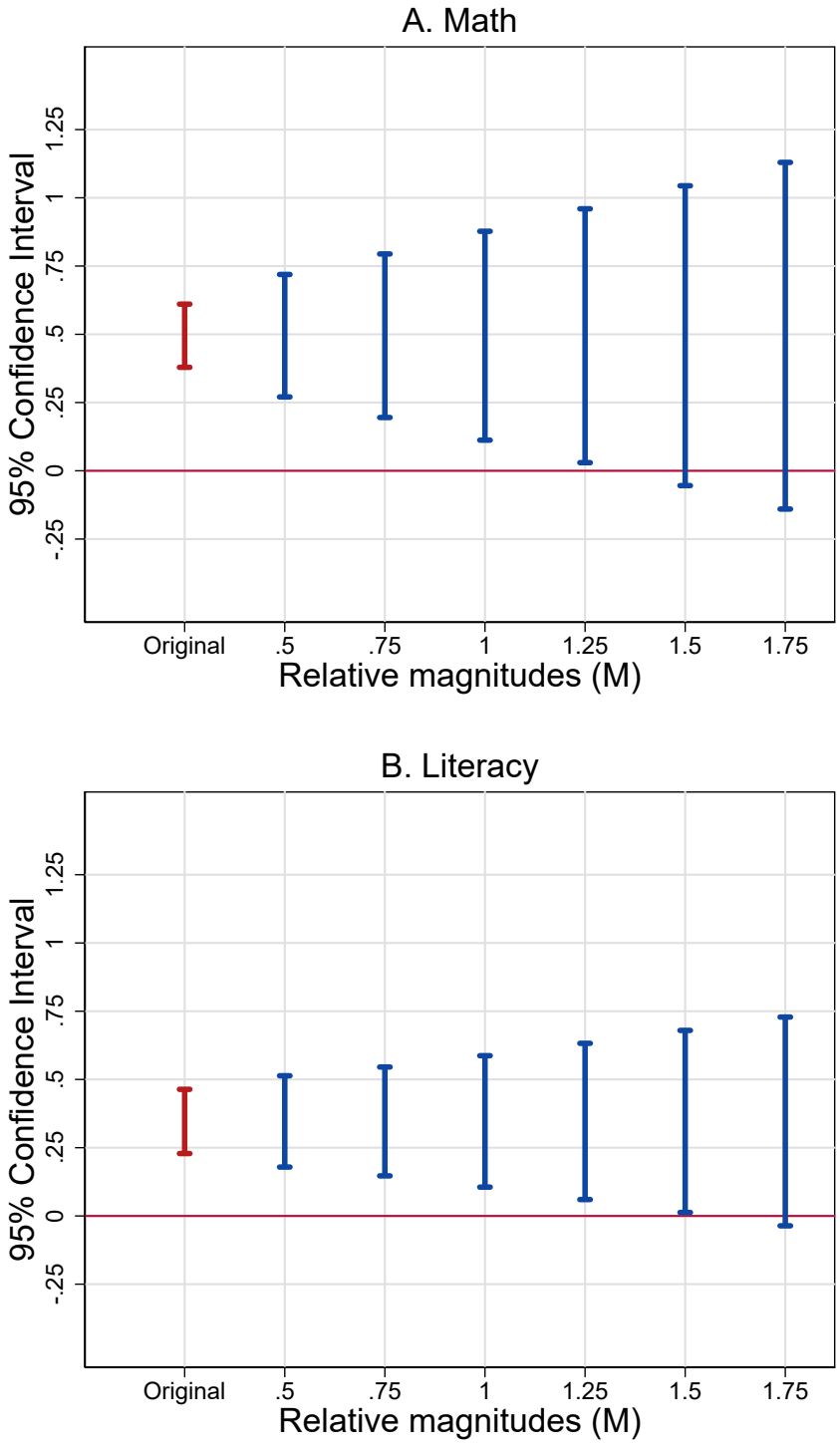


Figure 7: Upper and lower bounds on the post-treatment (lag) coefficient of period +4 using the honest difference-in-differences approach from Rambachan and Roth (2023). Notes: Relative magnitudes are the relative changes in the maximal (or worst) pre-treatment (lead) coefficient between consecutive years. Bar in red shows the original bounds on the lag coefficient of period +4. Each blue bar shows the bounds on the lag coefficient if the maximal (or worst) lead coefficient changes by different relative magnitudes (M).

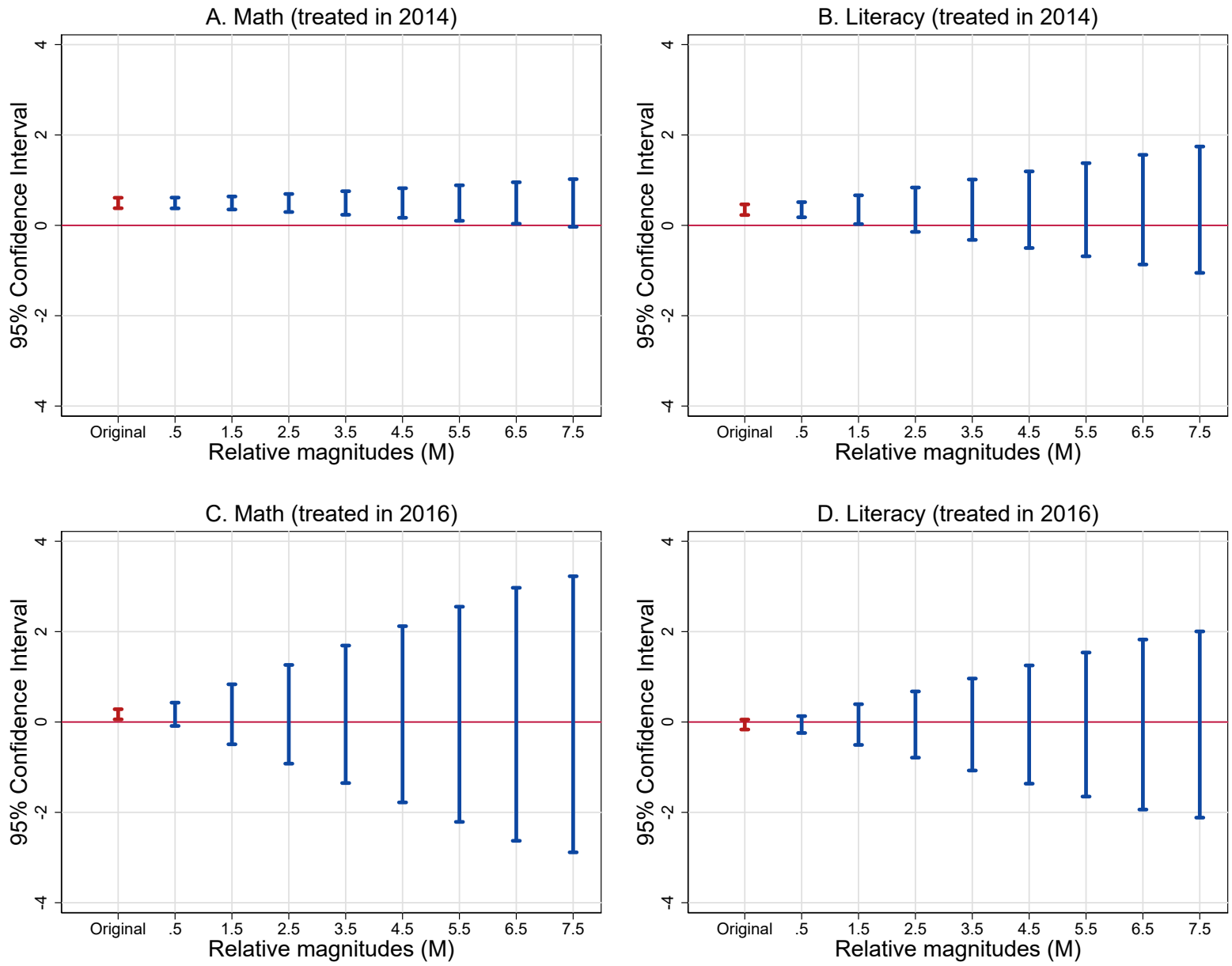


Figure 8: Upper and lower bounds on the post-treatment (lag) coefficient of each treatment group using the honest difference-in-differences relative magnitudes approach from Rambachan and Roth (2023). Notes: Relative magnitudes are the relative changes in the maximal (or worst) pre-treatment (lead) coefficient between consecutive years. Bar in red shows the original bounds on the lag coefficient of period +4 for districts treated in 2014, and period +2 for districts treated in 2016. Each blue bar shows the bounds on the lag coefficient if the maximal (or worst) lead coefficient changes by different relative magnitudes (M).

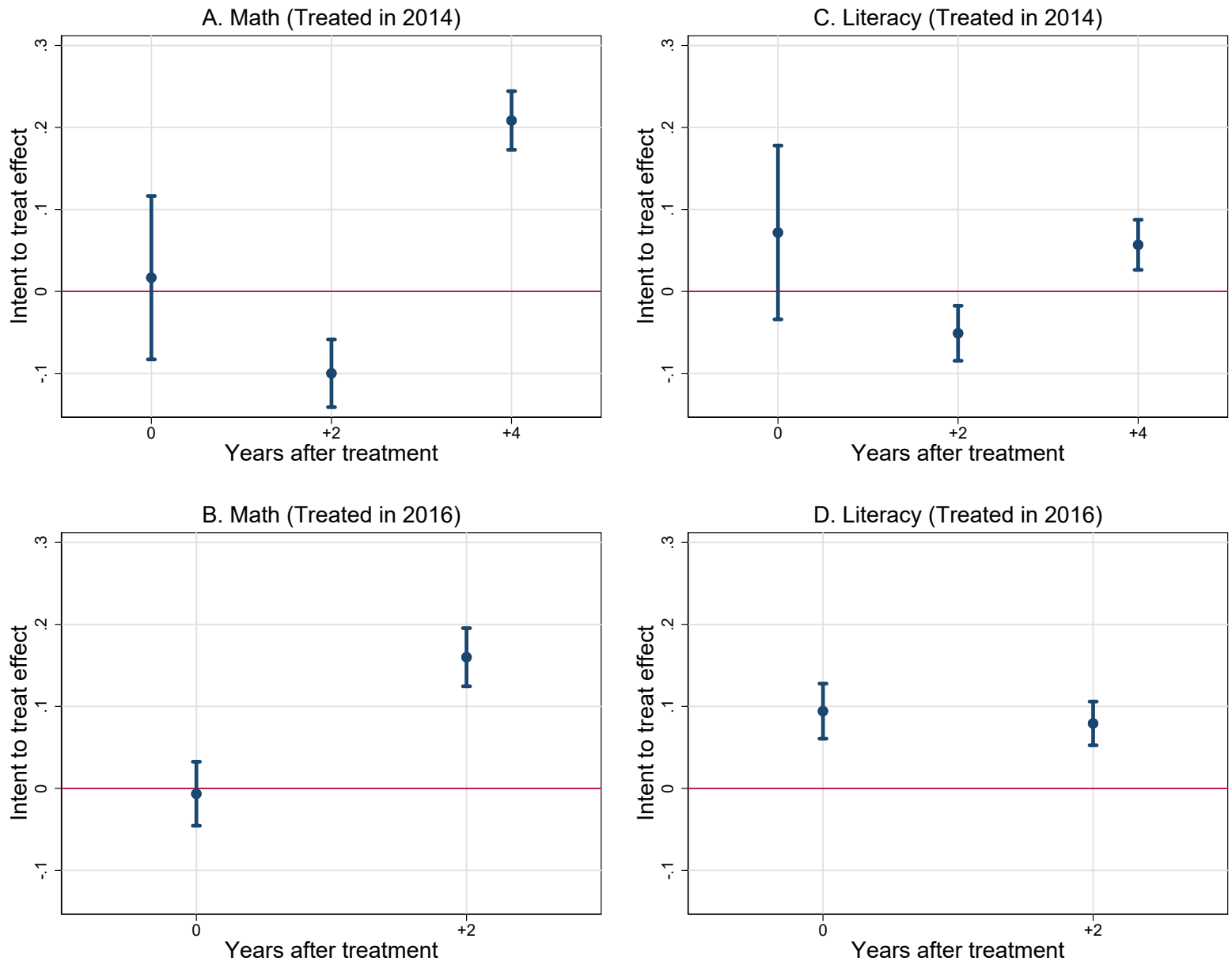


Figure 9: Math and literacy intent to treat effects of Clean India Mission using Wooldridge (2021). Notes: Post-treatment years for districts treated in 2014, are 2016 and 2018. Post-treatment years for districts treated in 2016, is 2018. Each bar represents the 95% confidence intervals and includes the point estimates.

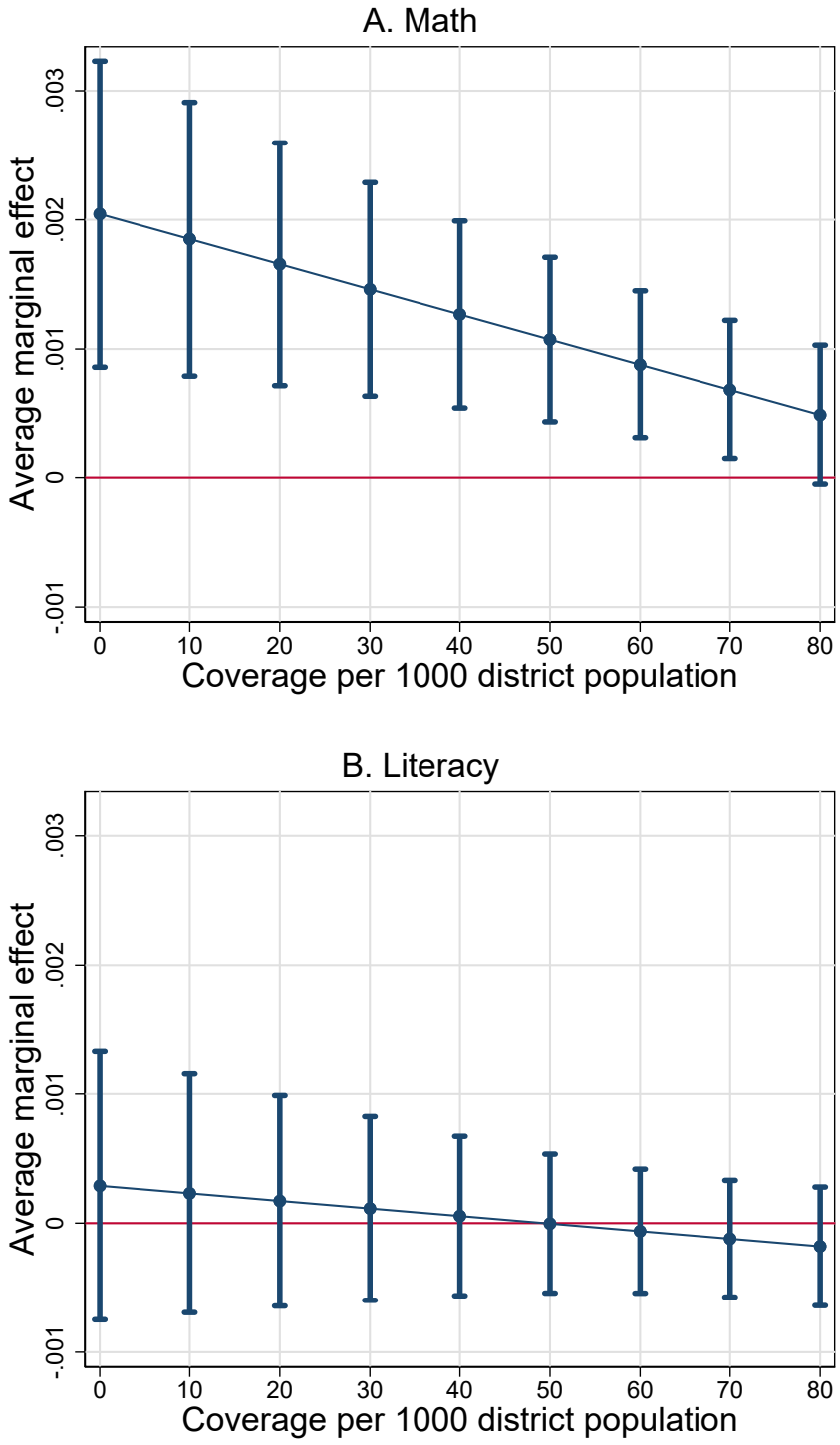


Figure 10: Average marginal effects of toilet construction through the Clean India Mission on children's math and literacy scores.

Table 1: Summary Statistics of base-year characteristics by year of coverage.

	Treated in 2014	Treated in 2016	Treated in 2018	Differences		
	N=331,069	N=13,933	N=1,917	2014 v. 2016	2014 v. 2018	2016 v. 2018
	(1)	(2)	(3)	(4)	(5)	(6)
Math	0.068	0.054	0.235	0.014	-0.167***	-0.181***
	(1.000)	(0.997)	(0.984)	(0.009)	(0.023)	(0.024)
Literacy	0.044	0.032	0.275	0.013	-0.231***	-0.244***
	(0.963)	(0.968)	(0.918)	(0.008)	(0.022)	(0.023)
Father attended school (1=Yes, 0=No)	0.724	0.856	0.744	-0.132***	-0.020*	0.112***
	(0.447)	(0.351)	(0.436)	(0.004)	(0.010)	(0.009)
Mother attended school (1=Yes, 0=No)	0.492	0.746	0.668	-0.254***	-0.176***	0.079***
	(0.500)	(0.435)	(0.471)	(0.004)	(0.011)	(0.011)
Male (1=Yes, 0=No)	0.530	0.524	0.562	0.006	-0.032***	-0.038***
	(0.499)	(0.499)	(0.496)	(0.004)	(0.011)	(0.012)
Age in years	10.402	10.535	10.605	-0.133***	-0.202***	-0.069
	(3.251)	(3.278)	(3.446)	(0.028)	(0.074)	(0.080)
Percent of villages with road	76.568	81.504	73.876	-4.936***	2.692***	7.628***
	(18.559)	(18.460)	(19.087)	(0.160)	(0.425)	(0.452)
Percent of villages with a post office	43.640	56.367	62.096	-12.727***	-18.457***	-5.729***
	(20.186)	(27.975)	(13.748)	(0.178)	(0.462)	(0.649)
Percent of villages with a health center	41.651	56.652	77.004	-15.001***	-35.353***	-20.352***
	(17.940)	(25.907)	(17.719)	(0.159)	(0.411)	(0.610)
Percent of school enrolled children	94.882	98.099	97.974	-3.217***	-3.092***	0.125**
	(4.200)	(1.972)	(2.302)	(0.036)	(0.096)	(0.049)
Percent of children in govt. school	70.181	61.840	84.735	8.341***	-14.554***	-22.895***
	(20.202)	(18.779)	(15.567)	(0.174)	(0.462)	(0.449)
Percent of households with a mobile phone	73.555	83.757	84.073	-10.202***	-10.518***	-0.316
	(16.196)	(13.424)	(6.616)	(0.139)	(0.370)	(0.312)
Percent of households with a TV	45.156	80.005	65.123	-34.849***	-19.967***	14.882***
	(25.459)	(18.850)	(29.137)	(0.218)	(0.584)	(0.496)
Percent of households with electricity	71.025	92.015	96.934	-20.990***	-25.909***	-4.919***
	(24.563)	(12.274)	(1.857)	(0.209)	(0.561)	(0.281)
Percent of households living in a thatched house	35.701	19.309	40.181	16.391***	-4.480***	-20.871***
	(24.131)	(19.686)	(19.874)	(0.207)	(0.552)	(0.480)
CIM coverage per 1000 district population	53.687	33.915	10.217	19.772***	43.470***	23.698***
	(35.959)	(25.992)	(2.650)	(0.308)	(0.821)	(0.594)

Notes: Average values of 2011 (base-year) characteristics in columns 1, 2, and 3. Differences in average values between the three treatment groups are shown in columns 4, 5, and 6. Standard deviations (in columns 1, 2, and 3) and standard errors (in columns 4, 5, and 6) are reported in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 2: Aggregated treatment effect estimates of Clean India Mission.

	(1) Math	(2) Literacy
Panel A: Intent to treat effect		
Weighted average (ITT)	0.166***	0.195***
	(0.0357)	(0.0364)
Panel B: Group specific effects		
Single parameter	0.166***	0.194***
	(0.0356)	(0.0362)
Treated in 2014	0.167***	0.197***
	(0.0360)	(0.0366)
Treated in 2016	0.104**	-0.0481
	(0.0409)	(0.0416)
Panel C: Calendar time effects		
Single parameter	0.168***	0.196***
	(0.0359)	(0.0366)
2014	-0.0360	0.0683**
	(0.0270)	(0.0267)
2016	0.0498	0.180***
	(0.0549)	(0.0565)
2018	0.491***	0.341***
	(0.0585)	(0.0594)
N	1836734	1836734

Notes: Aggregated impact of district-level toilet construction through the Clean India Mission on children's cognition using the doubly robust difference-in-differences approach from Callaway and Sant'Anna (2021). Panel A reports the weighted average of all available group-time intent to treat effects. Weights are based on the respective group sizes. Panel B summarizes treatment effects by the timing of the first year of coverage. These years are 2014 and 2016. All districts are treated by 2018, so I do have a treatment effect for 2018 districts. Panel C summarizes treatment effects by each post-treatment ASER survey year. These years are 2014, 2016, and 2018. Both models include all specified covriates. Standard errors clustered by district are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 3: Heterogeneity in the impact of Clean India Mission on children's cognition across gender and age groups.

	Math		Literacy	
	(1)	(2)	(3)	(4)
Panel A: Gender	Boys	Girls	Boys	Girls
Weighted average (ITT)	0.151***	0.178***	0.198***	0.185***
	(0.0466)	(0.0561)	(0.0478)	(0.0558)
N	941369	895365	941369	895365
	(1)	(2)	(3)	(4)
Panel B: Age groups	5 - 10 years	11 - 16 years	5 - 10 years	11 - 16 years
Weighted average (ITT)	0.196***	0.0718	0.258***	0.0382
	(0.0492)	(0.118)	(0.0528)	(0.122)
N	971854	864880	971854	864880

Notes: Aggregated impact of Clean India Mission using the doubly robust difference-in-differences approach from Callaway and Sant'Anna (2021). Boys and girls in Panel A, and younger (5 - 10 years) and older (11 - 16 years) children in Panel B. Covariates include whether the child's father and mother attended school, gender, age, age^2. District level covariates from 2011 are percent of villages with a road, a post office, a healthcare center, percent of school enrolled children, percent of children in government schools, percent of households with a TV, with electricity, and the percentage of households living in a thatched dwelling. Standard errors clustered by district are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 4: Group-time intent to treat effects of Clean India Mission on children's cognition.

	(1) Math	(2) Literacy
Panel A: Treated in 2014		
2011 v. 2012	0.00404	0.0709***
	(0.0251)	(0.0246)
2012 v. 2013	0.00549	-0.0459*
	(0.0260)	(0.0263)
2013 v. 2014	-0.0360	0.0683**
	(0.0270)	(0.0267)
2013 v. 2016	0.0499	0.182***
	(0.0555)	(0.0571)
2013 v. 2018	0.495***	0.346***
	(0.0591)	(0.0600)
Panel B: Treated in 2016		
2011 v. 2012	0.202**	0.139*
	(0.0874)	(0.0781)
2012 v. 2013	-0.0532	0.0144
	(0.0755)	(0.0739)
2013 v. 2014	-0.0826*	0.110**
	(0.0439)	(0.0443)
2014 v. 2016	0.0380	-0.0397
	(0.0458)	(0.0459)
2014 v. 2018	0.170***	-0.0565
	(0.0571)	(0.0563)
N	1836734	1836734

Notes: Group-time intent to treat effects using the doubly robust difference-in-differences approach from Callaway and Sant'Anna (2021). The treated group is composed of districts that receive Clean India Mission coverage in 2014 or 2016, and the control group is composed of districts that are not-yet-treated/covered by the program. Covariates include whether the child's father and mother attended school, gender, age, age^2. District level covariates from 2011 are percent of villages with a road, a post office, a healthcare center, percent of school enrolled children, percent of children in government schools, percent of households with a TV, with electricity, and the percentage of households living in a thatched dwelling. Standard errors clustered by district are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 5: Aggregated Event study.

	(1) Math	(2) Literacy
Pre treatment average	0.0699**	0.0551**
	(0.0303)	(0.0273)
Post treatment average	0.170***	0.198***
	(0.0362)	(0.0368)
-3	0.202**	0.139*
	(0.0874)	(0.0781)
-2	0.00338	0.0702***
	(0.0248)	(0.0243)
-1	0.00440	-0.0440*
	(0.0257)	(0.0260)
0	-0.0351	0.0670**
	(0.0267)	(0.0263)
+2	0.0514	0.179***
	(0.0548)	(0.0564)
+4	0.495***	0.346***
	(0.0591)	(0.0600)
N	1836734	1836734

Notes: Aggregated event study estimates using the doubly robust difference-in-differences approach from Callaway and Sant'Anna (2021). Base year for each lead year (-3, -2, and -1) is the respective previous year. Base year for each lag year (0, +2, and +4) is the year just before first program coverage. Coefficients on 'Pre treatment average' and 'Post treatment average' are the averages of lead and lag coefficients. Covariates include whether the child's father and mother attended school, gender, age, age^2. District level covariates from 2011 are percent of villages with a road, a post office, a healthcare center, percent of school enrolled children, percent of children in government schools, percent of households with a TV, with electricity, and the percentage of households living in a thatched dwelling. Standard errors clustered by district are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 6: Honest difference-in-differences.

	Math		Literacy	
	(1) Lower bound	(2) Upper bound	(3) Lower bound	(4) Upper bound
Original	0.379	0.611	0.229	0.464
M=0.50	0.271	0.719	0.179	0.514
M=0.75	0.195	0.794	0.147	0.546
M=1.00	0.112	0.877	0.106	0.587
M=1.25	0.030	0.960	0.060	0.632
M=1.50	-0.054	1.044	0.013	0.680
M=1.75	-0.140	1.130	-0.036	0.729

Notes: Upper and lower bounds of the post-treatment (lag) estimate of period +4 using the honest difference-in-differences relative magnitudes approach from Rambachan and Roth (2023). M represents the relative magnitudes, which is relative change in the maximal (or worst) pre-treatment (lead) coefficient between consecutive years. Values on the row titled 'Original' are the upper and lower bounds on the coefficient for period +4 from the aggregated event study design.

Table 7: Honest difference-in-differences for each treatment group.

	Math		Literacy	
	(1)	(2)	(3)	(4)
Panel A: Treated in 2014	Lower bound	Upper bound	Lower bound	Upper bound
Original	0.379	0.611	0.229	0.464
M=0.50	0.375	0.615	0.179	0.514
M=1.50	0.351	0.639	0.028	0.664
M=2.50	0.295	0.695	-0.145	0.837
M=3.50	0.234	0.756	-0.322	1.014
M=4.50	0.168	0.822	-0.502	1.194
M=5.50	0.101	0.887	-0.684	1.377
M=6.50	0.036	0.954	-0.867	1.560
M=7.50	-0.033	1.022	-1.052	1.744
	(1)	(2)	(3)	(4)
Panel B: Treated in 2016	Lower bound	Upper bound	Lower bound	Upper bound
Original	0.058	0.282	-0.167	0.054
M=0.50	-0.088	0.428	-0.243	0.130
M=1.50	-0.494	0.834	-0.507	0.394
M=2.50	-0.922	1.262	-0.789	0.676
M=3.50	-1.351	1.692	-1.074	0.961
M=4.50	-1.780	2.121	-1.365	1.252
M=5.50	-2.212	2.552	-1.650	1.537
M=6.50	-2.630	2.970	-1.938	1.825
M=7.50	-2.885	3.225	-2.117	2.004

Notes: Upper and lower bounds of the post-treatment (lag) estimate for each treatment group using the honest difference-in-differences relative magnitudes approach from Rambachan and Roth (2023). Panel A shows the results for districts treated in 2014. Panel B shows the results for districts treated in 2016. M represents the relative magnitudes, which is relative change in the maximal (or worst) pre-treatment (lead) coefficient between consecutive years. Values on the row titled 'Original' are the upper and lower bounds on the coefficient for period +4 for districts treated in 2014, and period +2 for districts treated in 2016 from the group-time or district-time intent to treat effect results.

Table 8: Potential mechanisms through which Clean India Mission affects children's cognition.

	(1) Toilet	(2) School Enr.	(3) Thatched
Weighted average (ITT)	0.0256	0.00763	-0.459***
	(0.0168)	(0.0113)	(0.0176)
N	1836734	1836734	1836734

Notes: Aggregated impact of Clean India Mission using the doubly robust difference-in-differences approach from Callaway and Sant'Anna (2021). Toilet is a binary variable for the household having its own toilet. School Enr. represents if the child is enrolled in a school or not. Thatched is a binary variable for type of dwelling for the household. It equals one for thatched housing, and equals zero for other dwelling type. Covariates include whether the child's father and mother attended school, gender, age, age^2. District level covariates from 2011 are percent of villages with a road, a post office, a healthcare center, percent of school enrolled children (except in column 2), percent of children in government schools, percent of households with a TV, with electricity, and the percentage of households living in a thatched dwelling (except in column 3). Standard errors clustered by district are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 9: Impact of Clean India Mission on alternative math and literacy test score measures.

	Raw score		Binary score		Age-standardized score	
	(1) Math	(2) Literacy	(3) Math	(4) Literacy	(5) Math	(6) Literacy
Weighted average (ITT)	0.200***	0.264***	0.0466**	0.0813***	0.202***	0.231***
	(0.0447)	(0.0510)	(0.0194)	(0.0194)	(0.0426)	(0.0445)
N	1836734	1836734	1836734	1836734	1836734	1836734

Notes: Aggregated impact of Clean India Mission using the doubly robust DD approach from Callaway and Sant'Anna (2021). In columns 1 and 2 the dependent variables are the raw math and literacy scores. In column 3, math equals 1 if the child has the highest math proficiency (which is level 4 - can do three by one digit division) and it equals zero for the other proficiency levels. In column 4, literacy equals 1 if the child has a level 4 proficiency (can read grade 2 text), and equals zero for all other levels. In columns 5 and 6, math and literacy scores are standardized using the age-specific mean and standard deviation measures in all the ASER rounds. Covariates include whether the child's father and mother attended school, gender, age, age^2. District level covariates from 2011 are percent of villages with a road, a post office, a healthcare center, percent of school enrolled children, percent of children in government schools, percent of households with a TV, with electricity, and the percentage of households living in a thatched dwelling. Standard errors clustered by district are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 10: Determinants of program implementation.

	(1) Coverage per 1000	(2) Treated in 2014	(3) Treated in 2014 or 2016
Percent of villages with road	0.0145	-0.000571	0.000269
	(0.0158)	(0.000532)	(0.000231)
Percent of villages with post office	0.00547	-0.000323	-0.000105
	(0.0286)	(0.000495)	(0.000153)
Percent of villages with health center	-0.0516	-0.000436	-0.0000391
	(0.0686)	(0.000684)	(0.000172)
Percent of fathers who attended school	-0.118*	0.00118*	0.000206
	(0.0581)	(0.000596)	(0.000308)
Percent of mothers who attended school	0.0987*	-0.000159	0.000645
	(0.0567)	(0.000830)	(0.000479)
Percent of school enrolled children	-0.116*	-0.00264	-0.00136
	(0.0617)	(0.00251)	(0.00209)
Percent of children in govt. school	0.0196	-0.0000298	0.0000405
	(0.0162)	(0.000397)	(0.000103)
Percent of households with mobile phone	-0.0639	-0.0000932	-0.0000405
	(0.0663)	(0.000349)	(0.000103)
Percent of households with TV	-0.0831	-0.000734	-0.000296
	(0.0786)	(0.000472)	(0.000219)
Percent of households with electricity	0.0927	0.000608*	0.0000884
	(0.0630)	(0.000346)	(0.000108)
Percent of households living in a thatched house	-0.0490	-0.000546**	-0.000253
	(0.0553)	(0.000244)	(0.000177)
Constant	23.36*	0.635***	1.082***
	(12.78)	(0.195)	(0.157)
N	1473	1473	1473
R^2	0.341	0.467	0.394

Notes: Associations between different pre-treatment (2011, 2012, and 2013) district level characteristics, and the respective dependent variable. In column 1, the dependent variable is the total number of toilets constructed (per 1000 population) through the Clean India Mission in its first year of coverage. In column 2, the dependent variable equals one for districts first covered by the program in 2014, and it equals zero for districts first covered in 2016 or 2018. In column 3, the dependent variable equals one for districts first covered in 2014 or 2016, and it equals zero for districts first covered in 2018. Both regressions include fixed effects for state, and year. Data are at the district level. Standard errors clustered by state are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 11: Impact of Clean India Mission on children's cognition using an alternative specification.

	(1) Math	(2) Literacy
2014 × 2014 × Treat	0.0168	0.0718
	(0.0507)	(0.0539)
2014 × 2016 × Treat	-0.0999***	-0.0511***
	(0.021)	(0.0171)
2014 × 2018 × Treat	0.2085***	0.0569***
	(0.0182)	(0.0156)
2016 × 2016 × Treat	-0.0066	0.0942***
	(0.0199)	(0.0171)
2016 × 2018 × Treat	0.1601***	0.0792***
	(0.018)	(0.0136)
Constant	-3.418***	-3.619***
	(0.2008)	(0.2551)
Agg. (Intent to treat effect)	0.0414*	0.0284
	(0.0213)	(0.0221)
N	1836734	1836734
R^2	0.443	0.467

Notes: Staggered difference-in-differences results using Wooldridge (2021). Both regressions include fixed effects for districts and survey years. Covariates include whether the child's father and mother attended school, gender, age, age^2. District level covariates from 2011 are percent of villages with a road, a post office, a healthcare center, percent of school enrolled children, percent of children in government schools, percent of households with a TV, with electricity, and the percentage of households living in a thatched dwelling. Standard errors clustered by district are reported in parentheses *** p<0.01, ** p<0.05, * p<0.1

Table 12: Impact of number of toilets constructed through the program on children's cognition.

	(1) Math	(2) Literacy
Coverage per 1000	0.00204***	0.000290
	(0.000603)	(0.000528)
(Coverage per 1000)^2	-0.00000972***	-0.00000293
	(0.00000363)	(0.00000326)
Constant	-4305.4***	-2877.3***
	(195.1)	(159.1)
N	1836734	1836734
R^2	0.444	0.468
Mean coverage per 1000	53.57	53.57
	(36.6)	(36.6)

Notes: Effect of Clean India Mission on children's math and literacy scores using a continuous treatment variable. Coverage per 1000 is the number of constructed toilets divided by district population multiplied by one thousand. Covariates include whether the child's father and mother attended school, gender, age, age^2. District level covariates from 2011 are percent of villages with a road, a post office, a healthcare center, percent of school enrolled children, percent of children in government schools, percent of households with a TV, with electricity, and the percentage of households living in a thatched dwelling. Both models include districts dummies, survey year dummies, and district specific time trends. Standard errors clustered by district are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Appendix

Table A.1: Aggregated treatment effect estimates of Clean India Mission (without covariates).

	(1) Math	(2) Literacy
Panel A: Intent to treat effect		
Weighted average (ITT)	0.0974***	0.0881***
	(0.0175)	(0.0179)
Panel B: Group specific effects		
Single parameter	0.0986***	0.0861***
	(0.0174)	(0.0178)
Treated in 2014	0.0949***	0.0921***
	(0.0177)	(0.0182)
Treated in 2016	0.180***	-0.0488*
	(0.0273)	(0.0277)
Panel C: Calendar time effects		
Single parameter	0.0971***	0.0877***
	(0.0174)	(0.0179)
2014	-0.00636	0.0256**
	(0.0115)	(0.0117)
2016	0.108***	0.154***
	(0.0284)	(0.0293)
2018	0.190***	0.0832***
	(0.0297)	(0.0299)
N	1836734	1836734

Notes: Covariates are not included in these regressions. Aggregated impact of district-level toilet construction through the Clean India Mission on children's cognition using the doubly robust difference-in-differences approach from Callaway and Sant'Anna (2021). Panel A reports the weighted average of all available group-time intent to treat effects. Weights are based on the respective group sizes. Panel B summarizes treatment effects by the timing of the first year of coverage. These years are 2014 and 2016. All districts are treated by 2018, so I do have a treatment effect for 2018 districts. Panel C summarizes treatment effects by each post-treatment ASER survey year. These years are 2014, 2016, and 2018. Standard errors clustered by district are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A.2: Group-time intent to treat effects of Clean India Mission on children's cognition (without covariates).

	(1) Math	(2) Literacy
Panel A: Treated in 2014		
2011 v. 2012	-0.0259**	-0.00614
	(0.0115)	(0.0113)
2012 v. 2013	0.0298***	0.0154
	(0.0113)	(0.0114)
2013 v. 2014	-0.00636	0.0256**
	(0.0115)	(0.0117)
2013 v. 2016	0.106***	0.161***
	(0.0289)	(0.0299)
2013 v. 2018	0.189***	0.0914***
	(0.0302)	(0.0305)
Panel A: Treated in 2016		
2011 v. 2012	0.176***	0.176***
	(0.0334)	(0.0325)
2012 v. 2013	-0.0112	0.0628**
	(0.0313)	(0.0319)
2013 v. 2014	-0.0228	0.212***
	(0.0299)	(0.0307)
2014 v. 2016	0.153***	0.00142
	(0.0316)	(0.0323)
2014 v. 2018	0.208***	-0.100***
	(0.0328)	(0.0329)
N	1836734	1836734

Notes: Group-time intent to treat effects using the doubly robust difference-in-differences approach from Callaway and Sant'Anna (2021). The treated group is composed of districts that receive program coverage in 2014 or 2016, and the control group is composed of districts that are not-yet-treated/covered by the program. Covariates are not included in these regressions. Standard errors clustered by district are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A.3: Aggregated Event study (without covariates).

	(1) Math	(2) Literacy
Pre treatment average	0.0595***	0.0655***
	(0.0115)	(0.0111)
Post treatment average	0.0999***	0.0887***
	(0.0177)	(0.0182)
-3	0.176***	0.176***
	(0.0334)	(0.0325)
-2	-0.0253**	-0.00313
	(0.0111)	(0.0109)
-1	0.0274**	0.0241**
	(0.0109)	(0.0110)
0	0.000346	0.0246**
	(0.0111)	(0.0113)
+2	0.111***	0.150***
	(0.0277)	(0.0286)
+4	0.189***	0.0914***
	(0.0302)	(0.0305)
N	1836734	1836734

Notes: Aggregated event study estimates using the doubly robust difference-in-differences approach from Callaway and Sant'Anna (2021). Base year for each lead year (-3, -2, and -1) is the respective previous year. Base year for each lag year (0, +2, and +4) is the year just before first program coverage. Coefficients on 'Pre treatment average' and 'Post treatment average' are the averages of lead and lag coefficients. Covariates are excluded from these regressions. Standard errors clustered by district are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A.4: Honest difference-in-differences (Smoothness bounds).

	Math		Literacy	
	(1) Lower bound	(2) Upper bound	(3) Lower bound	(4) Upper bound
Original	0.379	0.611	0.229	0.464
N=0.05	0.343	0.656	0.144	0.461
N=0.15	0.243	0.755	0.044	0.560
N=0.25	0.143	0.855	-0.056	0.660
N=0.35	0.043	0.955	-0.156	0.760
N=0.45	-0.057	1.055	-0.256	0.860

Notes: Upper and lower bounds of the post-treatment (lag) estimate of period +4 using the honest difference-in-differences smoothness bounds approach from Rambachan and Roth (2023). N represents the absolute value by which post-treatment differential trends can differ from the maximal pre-treatment differential trends. Formally, this is represented as $Post_{trends} = N \pm Pre_{trends}$. N=0 implies that differential trends are linear, i.e., pre-existing cognition trends between children in treatment and control districts evolve smoothly over time. Other values of N that I use are 0.05, 0.15, 0.25, 0.35, and 0.45. Values on the row titled 'Original' are the upper and lower bounds on the coefficient for period +4 from the event study design.

Table A.5: Honest difference-in-differences or each treatment group (Smoothness bounds).

	Math		Literacy	
	(1)	(2)	(3)	(4)
Panel A: Treated in 2014	Lower bound	Upper bound	Lower bound	Upper bound
Original	0.379	0.611	0.229	0.464
N=0.05	0.344	0.657	0.142	0.458
N=0.15	0.244	0.756	0.043	0.558
N=0.25	0.144	0.856	-0.057	0.658
N=0.35	0.044	0.956	-0.157	0.758
N=0.45	-0.056	1.056	-0.257	0.858
	(1)	(2)	(3)	(4)
Panel B: Treated in 2016	Lower bound	Upper bound	Lower bound	Upper bound
Original	0.058	0.282	-0.167	0.054
N=0.05	-0.064	0.273	-0.137	0.191
N=0.15	-0.183	0.358	-0.212	0.319
N=0.25	-0.283	0.458	-0.312	0.419
N=0.35	-0.383	0.558	-0.412	0.519
N=0.45	-0.483	0.658	-0.512	0.619

Notes: Upper and lower bounds of the post-treatment (lag) estimate of period +4 using the honest difference-in-differences approach from Rambachan and Roth (2023). N represents the absolute change in the largest pre-treatment (lead) coefficient between consecutive years. Values on the row titled 'Original' are the upper and lower bounds on the coefficient for period +4 for districts treated in 2014, and period +2 for districts treated in 2016 from the group-time or district-time intent to treat effect results.

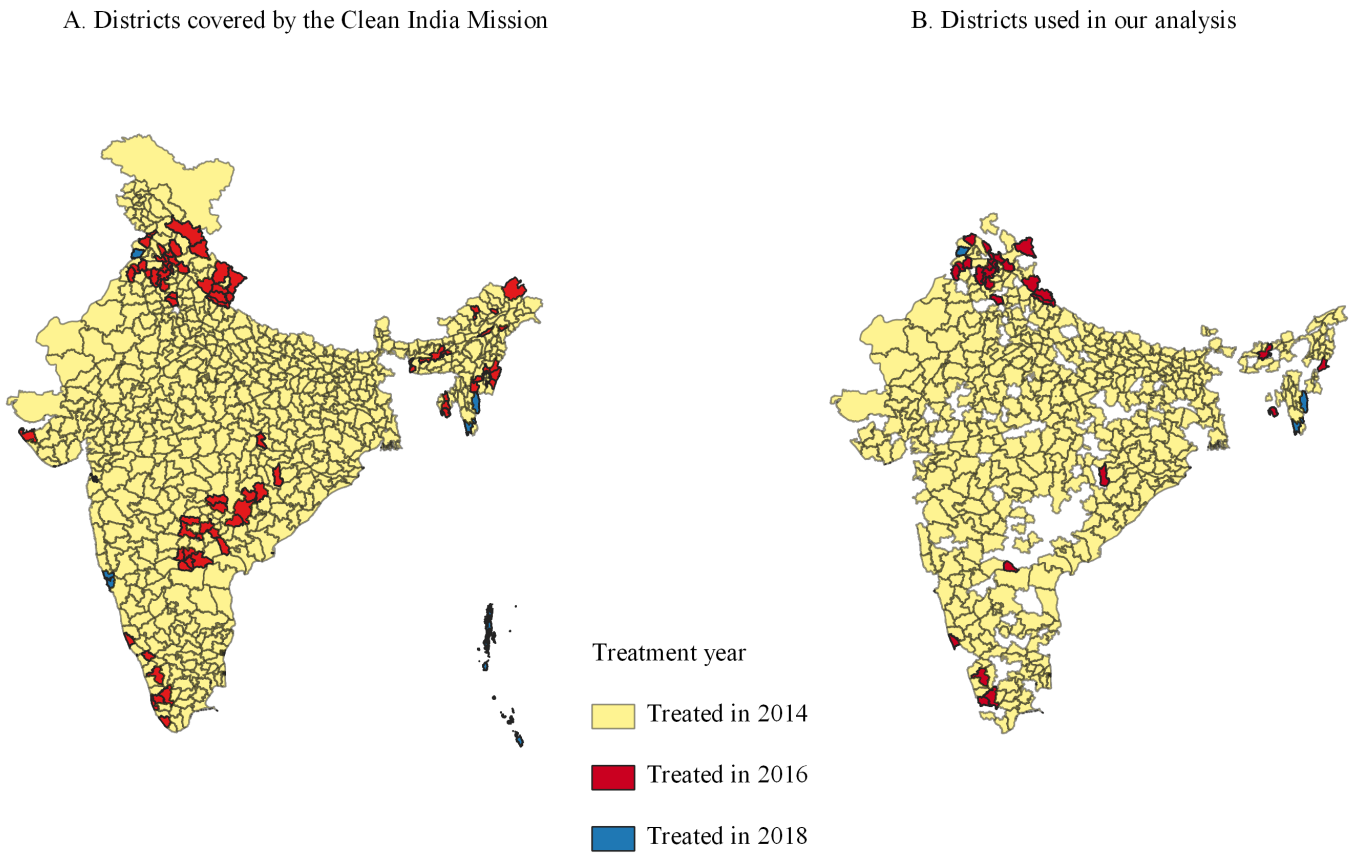


Figure A.1: First year of treatment for each district. Notes: Panel A shows all districts covered by the Clean India Mission from the Indian Government data. Panel B shows the districts used in this paper.

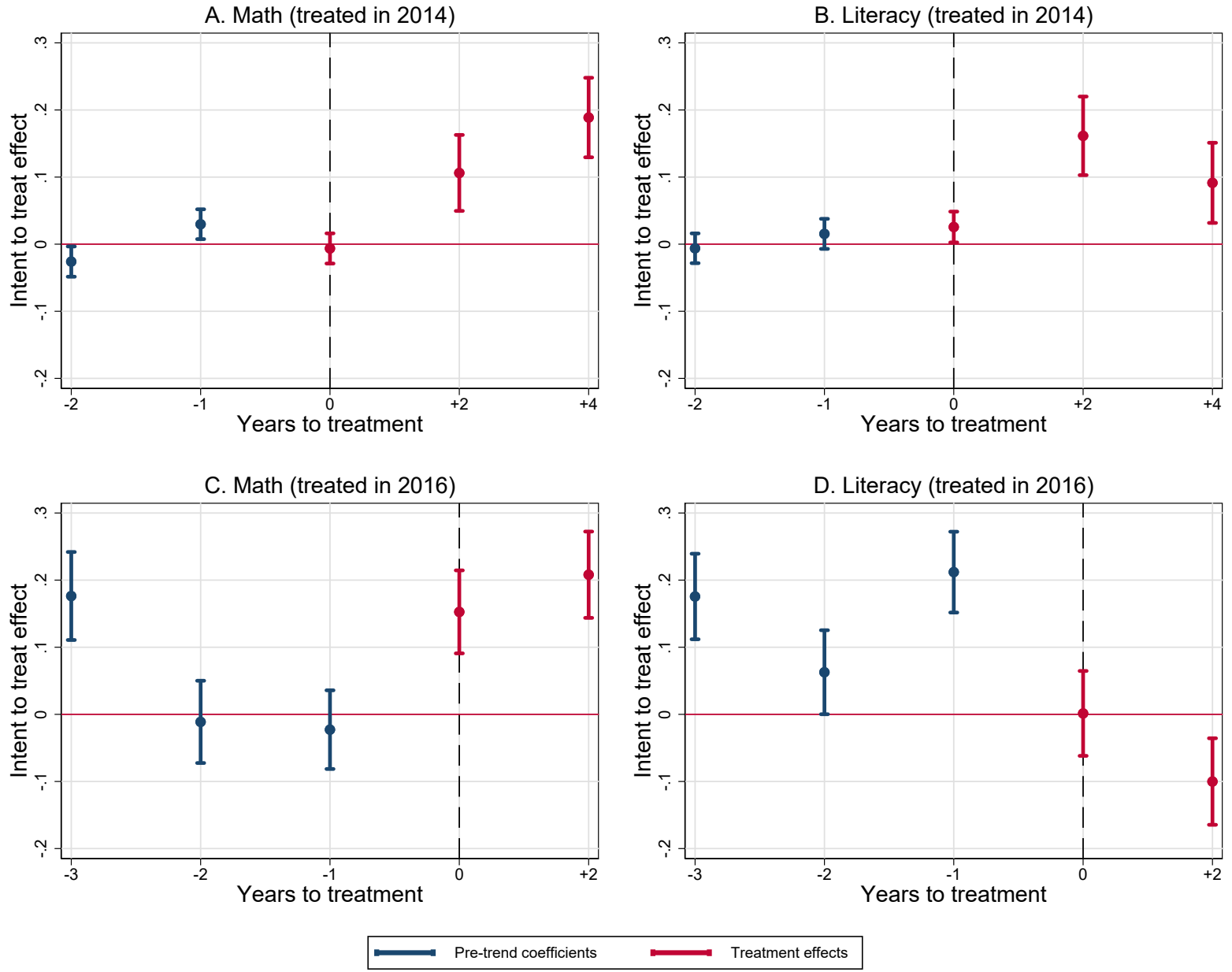


Figure A.2: Math and literacy group-time intent to treat effects of Clean India Mission using Callaway and Sant'Anna (2021). Notes: Control variables are not included in these results. Lag year coefficients are shown in red. Base year for each lag coefficient is the year just prior to first treatment. Lead coefficients are shown in blue, and the base year for each lead coefficient is the respective previous year. Parallel trends holds if all pre-treatment coefficients (lead coefficients) are statistically similar to zero. Each bar indicates the 95% confidence intervals and includes the point estimates.

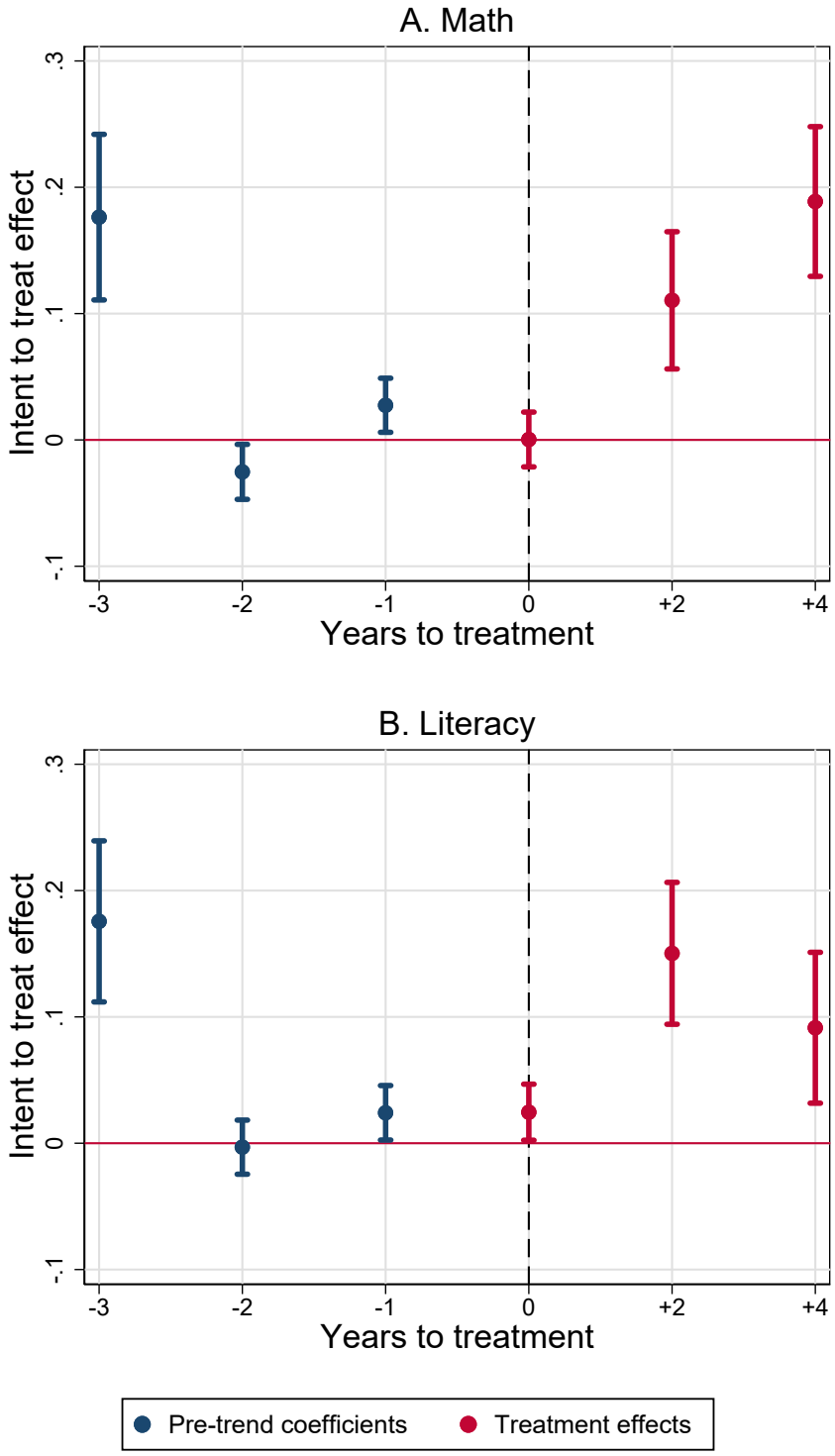


Figure A.3: Event study using the doubly robust difference-in-differences approach by Callaway and Sant'Anna (2021). Notes: Lag year coefficients are shown in red. Base year for each lag coefficient is the year just prior to first treatment. Lead coefficients are shown in blue, and the base year for each lead coefficient is the respective previous year. Parallel trends holds if all pre-treatment coefficients (lead coefficients) are statistically similar to zero. Each bar indicates the 95% confidence intervals and includes the point estimates. Covariates are excluded from these results.

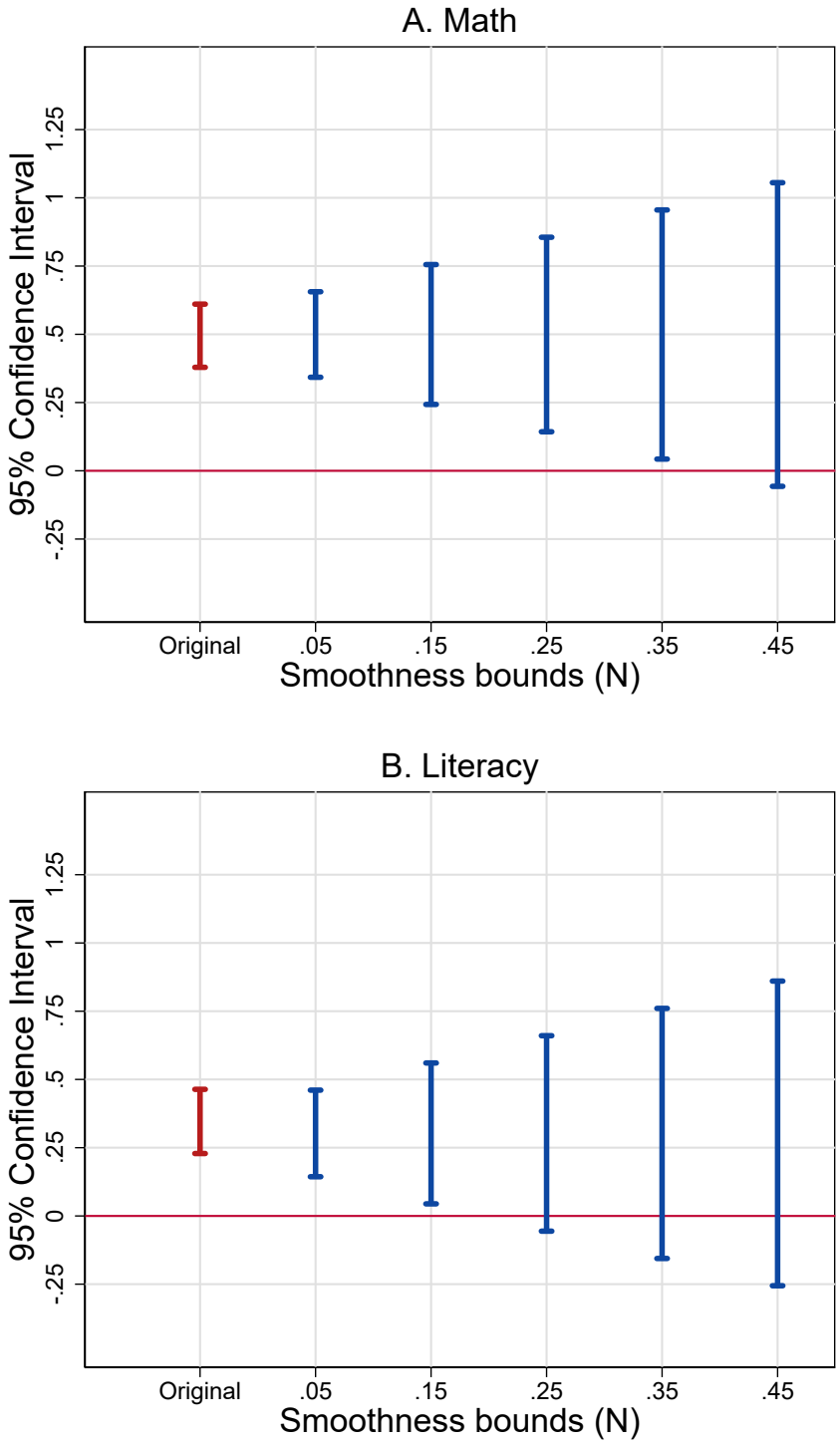


Figure A.4: Upper and lower bounds on the post-treatment (lag) coefficient of period +4 using the honest difference-in-differences smoothness bounds approach from Rambachan and Roth (2023). Notes: Bar in red shows the original bounds for the lag coefficient of period +4. Each blue bar shows the bounds on the lag coefficient if the largest lead coefficient changes by different absolute values (N). The breakdown point of the program's math score impact at period +4, is 0.45 SD. The literacy score breakdown point is 0.25 SD with confidence intervals of -0.056 SD, and +0.660 SD.

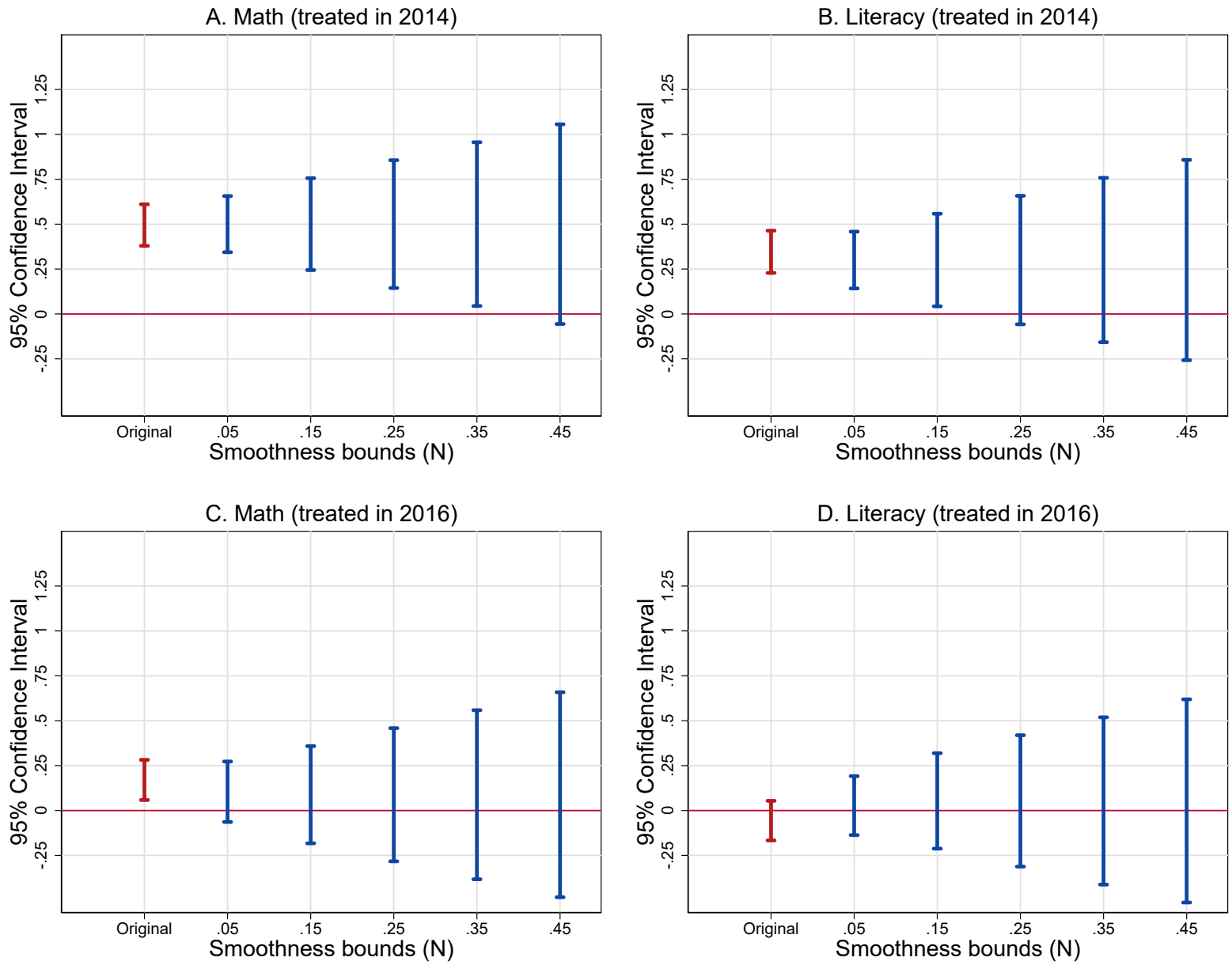


Figure A.5: Upper and lower bounds on the post-treatment (lag) coefficient using the honest difference-in-differences smoothness bounds approach from Rambachan and Roth (2023). Notes: Smoothness bounds represent the absolute changes in the largest pre-treatment (lead) coefficient between consecutive years. Bar in red shows the original bounds for the lag coefficient of period +4 for districts treated in 2014, and period +2 for districts treated in 2018. Each blue bar shows the bounds on the lag coefficient if the largest lead coefficient changes by different absolute values (N).