Child Cognition Benefits of the Clean India Mission Sanitation Program

Abhradeep Karmakar* Kira Villa†

(Job market paper)

Click here for the most updated version

Abstract

This work evaluates the impact of the Clean India Mission sanitation program on children's cognition. Using survey data from the ASER, and program coverage data from the Indian government, we utilize the district level roll-out of the program to examine changes in children's math and literacy (reading) test-scores. We find positive and significant improvements in math scores, but not in literacy scores. Results are similar across genders, and two different age-groups. Our findings violate the parallel trends assumption. To address this, we examine relative changes in differential pre-trends for an unbiased treatment effect estimate. Robustness checks point to dynamic improvements in math scores, and also show that math score improvements are not driven by co-existing village-level changes. Multiple pathways, through which the program can affect cognition, are recognized. To address non-random program roll-out, we test if pre-treatment village, district, and household-level characteristics predict program coverage.

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. Visiting Assistant Professor, Department of Economics, St. Lawrence University. Email - akarmakar@stlawu.edu

†Associate Professor, Department of Economics, University of New Mexico. Email - kmvilla@unm.edu

We thank Matías Fontenla, David Stifel, Melissa Binder, Tania Barham, Gautam Hazarika and participants of the Southern Economic Association (SEA) Conference 2022, and 2023, and the North East Universities Development Consortium (NEUDC) Conference 2023 for their feedback on this work.

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). Policies and programs aimed at improving sanitation access and reducing open defecation 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 also 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 Villa (2017).

In this paper, we evaluate the impact of the Clean India Mission (*Swachh Bharat Mission*, 2014) on children's cognition in rural areas. This program was implemented by the Indian Government in 2014 with the aim of reducing open defecation, improving toilet-access, household and community sanitation infrastructure, and overall living conditions across the country. For achieving these goals, attributes of the program include the construction of household and community-owned toilets, sewerage facilities, waste collection systems, and the elimination of manual scavenging (cleaning and disposal of waste from sewage pits). Inadequate sanitation infrastructure and the associated pollution can impact children's cognition 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 subsequently can cause increased illness incidence, which can reduce school participation and cognitive outcomes of children. 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), and Petri et al. (2008). Therefore, despite not specifically targeting cognition, household sanitation programs, such as the Clean India Mission (CIM), can impact children's cognitive performance by lowering the incidence of diseases, improving school participation and overall household quality.

To estimate the relationship between the Clean India Mission (CIM) and children's math and literacy scores, we combine data from seven rounds of the Annual Status of Education Report (2010, 2011, 2012, 2013, 2014, 2016 and 2018) and three rounds of district level coverage of the program (2014, 2015, 2017). We employ both a standard two way fixed effects (TWFE), and a continuous treatment effects empirical strategy, by exploiting the district-level differences in the program's coverage. Results indicate that CIM causes significant improvements in children's math scores for those in treated districts relative to control districts. Findings from the continuous coverage strategy show that an increase in toilets constructed through the program, per 1000 district population, leads to significant improvements in math scores. The program's impact on literacy scores is positive, but noisy and not consistently statistically significant. We also show that children in early adoption districts report significantly improved test-scores than those in the late/non adoption districts. Sub-group analyses reveal that both boys and girls, and younger and older children experience positive and significant math score improvements due to their exposure to the program. Event study results show that prior to the implementation of CIM, children in treatment districts report significantly different trends in math scores than those in the control districts. But in periods following the implementation, these differential trends are positive and improving over time. Event study results for literacy scores are not statistically significant in both the pre- and post-treatment periods.

Our findings are robust to multiple checks. To support the event study results, we adopt the honest difference-in-differences approach developed in Rambachan and Roth (2023). Through this approach, we account for pre-existing differential trends that potentially lead to biased treatment effects of CIM on children's cognition. We also adopt a novel empirical strategy to show that the quasi effects of toilet construction, prior to actual treatment (before 2014), on math scores are mostly negative, or statistically insignificant.

To exploit the staggered nature of implementation of the program, we adopt the method developed in Wooldridge (2021) and recover significantly improving treatment effects on math scores of children in the earliest treatment districts. These dynamic improvements do not hold for literacy scores. We show that our results are not driven by co-existing government programs that improve village-level infrastructure. As potential mechanisms, we find that an increase in the program's coverage (toilet construction in district) leads to positive changes in children's school attendance, households' toilet ownership, and a decline in the probability of families living in a thatched house. We also address concerns with non-random implementation of the program by showing that district-level coverage of CIM is not associated with multiple pre-treatment village, district, and household-level characteristics. Lastly, we focus on our dependent variables, and use raw and binary math and literacy scores as children's cognition outcomes, as the alternatives for standardized 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; Ortiz-Correa et al., 2016; Adukia, 2017; 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, we 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 credible evidence on the relationship between school and community infrastructure and children's health and nutrition, the effect of sanitation interventions on cognition is a relatively sparse area of study. Secondly, we add to an even smaller group of papers 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. 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. We also address the role of pre-existing programs that could influence the treatment effects we estimate for CIM.

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 heavily emphasizing on eradicating open defecation, improving sanitation access and water supply, public health, waste management, and household living conditions, evidence suggests that the effect of these programs on human capital 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).

Building on these existing programs, the Clean India Mission (CIM) was launched in October 2014 by the newly elected central government. The program was introduced in two phases. Objectives of Phase 1, which functioned 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.

The broader goal of the program 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 Clean India Mission (CIM) focuses on sustaining the gains of phase 1. Phase 2 operates between 2020 and 2025.

Implementation of CIM is monitored by two different government agencies – the Ministry of

Drinking Water and Sanitation in rural areas,[1] and the Ministry of Housing and Urban Affairs in urban areas.[2] 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 CIM 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.[3] 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 typically 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; Hammer and Spears, 2016; Hueso and Bell, 2013; Kumar, 2017). A small selection of studies focuses on the effect of CIM on different factors. Dandabathula et al. (2019) show that CIM 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 CIM. 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.

2.2 Literature review

Broad evidence demonstrates that sanitation infrastructure improves child health and nutrition outcomes (Augsburg and Rodriguez-Lesmes, 2018; Bartram et al., 2005; Kumar and Vollmer, 2013; Mara et al., 2010; Pickering et al., 2015; Jasper et al., 2012; Spears, 2013). There also exists a well-

---

[1] In rural areas, CIM is referred to as "Swachh Bharat Mission–Gramin." Gramin translates into rural.

[2] In urban areas, CIM is referred to as Swachh Bharat Mission – Urban

[3] 2019 CIM report

established body of literature identifying factors like waste exposure, and air pollution (Bharadwaj et al., 2017; Saenz et al., 2018; Dix-Cooper et al., 2012), lead, arsenic, and toxin exposure through contaminated water, all of which are direct outcomes of inadequate sanitation access (Aizer et al., 2018; Ruff et al., 1996; Rosado et al., 2007; Liu and Lewis, 2014), causing 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. Similarly, research like Cameron et al. (2021), Cronk et al. (2015), Greenberg et al. (1999), Ortiz-Correa et al. (2016), and Grantham-McGregor et al. (1999), report associations between cognitive ability or school attendance, and sanitation facilities in children's neighborhood. Their findings discuss that in areas where communities lack proper sanitation infrastructure, children are more prone to water-borne 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), and Suryadarma et al. (2006) study school-level programs focused on school infrastructural facilities (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, we are the first to examine these cognition effects in India.

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 approximately cover 600,000 children aged 3-16 years, across 20 randomly selected rural households in 20-30 villages in around 600 rural districts in India. The surveys collect information on household size, parental education, household assets and infrastructure, gender, and schooling status of children, and also test those aged 5–16 on basic literacy (reading) and math proficiency. In our study, we use data from seven rounds (2010, 2011, 2012, 2013, 2014, 2016, and 2018).

The literacy and math achievement tests, employed by ASER, score learning proficiency on a scale from 0–4. 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 1 presents the distribution of raw literacy and math scores for our sample children. We find that children, on average, have higher math than literacy abilities except for the highest proficiency level (score 4). Only 26.6 percent of them can perform a three-by-one-digit division, whereas 43.3 percent can read grade 2 texts. The figure also shows that 13 percent of the children cannot read, and 11.2 percent cannot do basic arithmetic.[4]

Table 1 reports summary statistics of the variables that we use. Math and literacy are our dependent variables. These represent the standard deviation (SD) differences between the raw math and literacy (reading) scores of a child and the mean of raw math and literacy scores of children in the district. Mean and standard deviations are calculated for each district based on the district-specific mean and standard deviation over the entire duration of the ASER surveys (2010 – 2018). We find that the mean values for both math and literacy are positive for 2010 and 2011, but they are negative for the other rounds. These estimates are -0.078 SD, -0.071 SD and -0.041 SD for 2014, 2016 and 2018, which indicates that between 2010 and 2014, mean math scores have declined, but from 2014 onwards they have improved.[5] Mean literacy scores are nearly equal between 2012 and 2014, and they have gradually improved from 2014 to 2018. Panel A of Figure 2 plots these trends in the standardized scores for math and literacy, showing small improvements from 2014 to 2018.

Table 1 reports that the sample is adequately represented by both boys and girls, who have a mean age of roughly 10 years, and approximately 97 percent of whom currently attend school. Parental school attendance, indicating the percentages of children whose parents have completed at least one year of formal schooling, is around 70 percent for fathers and 55 percent for mothers. We find that approximately 25 percent to 30 percent of the sample children live in a thatched house. This percentage has declined from 2012 onwards, but they are lowest in 2016 and 2018. These values are 30 percent and 26 percent respectively. Toilet in house is a binary indicator of whether the child's household has its own toilet. Table 1 and panel B of Figure 2 show that less than 45 percent of the sample children have a household toilet during the rounds of 2010 – 2014, but this estimate increases sharply to 54 percent and 72 percent in 2016 and 2018. Table 1 also presents

---

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

[5] Shah and Steinberg (2019) discuss that this decline in test-scores (between 2010 and 2014) is potentially associated with the 2009 Right to Education Act in India. This nationwide school-level legislation led to increases in classroom size, large influxes of lower ability student in private schools, and the universal promotion of students in primary schools. All these changes can reduce children's learning achievements.

data on households' access to different village-level facilities. These include a primary health center, a bank, and a post office. Access to a primary health center is around 42 percent across all the ASER rounds. The proportion of sample children living in a village with a banking facility has slowly grown from 22 percent to 29 percent between 2010 and 2018, but the percentage having a post office in the village has declined slightly between these years.

Data on the annual coverage of the Clean India Mission (CIM) come from the web portal of 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 CIM in each of the post-CIM years, 2014, 2015, 2016 – 2021.[6] ASER does not provide the block (smaller sampling unit than district) names, so we match the CIM coverage data for 2014, 2015, and 2017, with the ASER rounds of 2014, 2016, and 2018 at the district level. This means ASER (2014) is matched with the current CIM coverage in 2014, ASER (2016) with the lagged CIM coverage in 2015, and ASER (2018) with the lagged CIM coverage for 2017. For the 2016 and 2018 ASER rounds, we use the lagged values (coverage in previous year) as this ensures that children have experienced at least one full year of exposure to the program during these years. Since CIM was first implemented in 2014, for ASER (2014) we are only able to match the current coverage data (ASER, 2014 and CIM, 2014).

The repeated cross-sectional ASER data, upon merging with the CIM coverage information, include a total of 588 districts between 2010 and 2018. Among these, 519 districts (2,283,175 observations) first appear (or first treated) in CIM (2014), followed by an additional 58 districts (147,400 observations) that are first treated in CIM (2015). Late adopters, comprising 9 new districts (20,434 observations), receive initial exposure to CIM in 2017. Across all the ASER rounds, we only have 2 districts (5,778 observations) that are never exposed to the program. Because of the repeated cross-sectional nature of the ASER surveys, across all the matched ASER-CIM coverage data (2010 – 2018) we identify 60 districts in 2010, 15 districts in 2011, 2 districts in 2012, 22 districts in 2016, and 2 districts in 2018, that are not consistently (i.e., observed with interruption

---

[6] Data for 2013 (prior to CIM) are available on the number of above/below poverty line rural households in each block (or district) that have/do not have their own toilet, and the total number of rural households in each district.

or attrition) available in all the years (2010 through 2018). These districts are included in our analysis.

Using the matched (ASER and CIM) data, Table 1 shows that on average, 8438, 19232, and 48984 new household toilets were constructed in the rural areas of each district in 2014, 2016, and 2018, respectively. CIM coverage data for the previous (lagged) year show that 9082 and 32745 household toilets are constructed in each district in 2015 and 2017, respectively. Using data from the 2011 census, we convert the CIM coverage information into the coverage rate per 1000 rural population in the district. Table 1 shows that, for 2014, 2016, and 2018, on average, approximately 6.73, 15.47, and 34.82 toilets are constructed for every 1000 people in rural regions of each district. These average coverage rates are small, but they are increasing overtime. For 2015 and 2017 these values are 6.97 and 25.39 per 1000, respectively.

The Clean India Mission (CIM) website presents information on the total number of households in the district in 2013 (i.e., prior to CIM). Using these data, we calculate the percentage of rural households in the district that have their own toilet. This percentage increases sharply from 38 percent in 2014, to 49 percent in 2016, and 80 percent in 2018.[7] Using the population census data from 2011, we find that the percentage of rural population with own household toilet is relatively small in each year, but it approximately doubles from 7.6 percent in 2014, to 15.9 percent in 2018.

Figure 3 shows the percentage of rural households in a district that have their own toilets between 2013 and 2018. It is noteworthy that most of the districts move from a lighter (less than 25 percent of the households) to a darker (more than 75 percent) shade within this period. This nature of household-level toilet access can also be established from the ASER surveys, as shown in panel B of Figure 2. From this figure, we find a substantial surge in the percentage of sample children with a household toilet facility from 2014 onwards.

Appendix Table A.1 presents the summary statistics of all the variables when we construct a balanced panel. In this balanced panel, we omit districts that could not be matched across all

---

[7] These percentages are slightly smaller than those reported in Table 1 of Hossain et al. (2022) This is because, our data only include the districts that we could match across the ASER surveys and the CIM coverage website. Another difference is that our percentage values are aggregated at the district level, whereas Hossain et al. (2022) present coverage information for each state in the table.

the ASER and CIM rounds. These omitted districts include districts both observed with interruption and attrition in the ASER surveys. In Table A.2, we show that average characteristics of the full sample of districts in column 1 do not differ from the balanced sample (column 2) by large amounts. These differences are shown in column 4. Similarly, in column 5, we find small differences in most of the average characteristics between the full sample and the omitted sample of districts. But we also find that children in districts that are omitted from some of the ASER surveys report a higher percentage of having a toilet facility at home. Greater access to household toilets, and the slightly larger percentages (less than 6 percentage points) of all the characteristics, except age, math and literacy, indicate that both the Clean India Mission and the ASER surveys primarily cover poor rural areas in districts. This is because households in poor rural regions/areas are more likely to live in a thatched house, not have proper access to a household toilet, or village level services (e.g., bank, post-office, and health-care facility), and children are less likely to attend school (Ravallion and Datt, 2002; Datt and Ravallion, 2002).[8]

4 Empirical Strategy

We use a two way fixed effects (TWFE) and a continuous treatment effect model to evaluate the relationship between CIM and children's cognition. Using both strategies we estimate the heterogeneity in the impact of the program across children's age and gender. These results are shown in section 4.3. Section 4.4 studies pre-treatment trends and reports the 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 developed in Rambachan and Roth (2023). Using the continuous coverage information, we estimate the quasi-treatment effect of the program in pre-treatment period. We adopt the staggered difference-in-differences method proposed in Wooldridge (2021) to estimate dynamic treatment effects of the program. We show that children's cognition benefits are not driven by village-level infrastructural changes resulting from

---

[8] In the regressions we use the full sample. Main results using the balanced sample are shown in the appendix.

co-existing government programs, and also identify multiple mechanisms through which the program can affect cognition. To address the non-random roll-out of the program, we test for the association between pre-treatment village, district, and household-level characteristics, and the program's district-level coverage. Lastly, we estimate the effect of the program using alternative cognition outcomes. For this test, we replace our standardized outcome variables with binary and raw math and literacy scores.

4.1 Two way fixed effects

Changes in standardized math and literacy scores due to the implementation of CIM are estimated using equation (1):

$$Z_{ibdt} = \alpha_0 + \alpha_1.(Post \times Treat)_{dt} + \alpha_2.X_i + \theta_d + \theta_t + \theta_b + \theta_d \times t + \varepsilon_{ibdt} \tag{1}$$

where, Z_{ibdt} is the standardized math or literacy score for a child i in district d time t, and born in year b. $Post$ takes a value of one for the years 2014, 2016, and 2018, and a value of zero for years prior to 2014. $Treat_{dt}$ measures treatment status and we measure it two ways.

First, we estimate equation (1) with $Treat_{dt}$ as a binary indicator of whether a district has been exposed to CIM in any of the post-treatment years, i.e., whether a toilet (latrine) is constructed in the district by CIM in 2014, 2015, or 2017.[9] Our control group includes districts never-exposed to the program, and also districts that first implemented the program in 2017 (i.e., the late-adoption districts). This leaves us with 2,451,009 treated and 5,778 control children if we compare CIM exposed (by 2014, 2016, and 2018) with never-exposed, and 2,430,575 and 26,212 children if we compare early-adopters (2014, and 2016) with late (2018) and never-adopters. Early adopters sustained their first CIM exposure in 2014, and 2015, and late adopters sustained their first exposure in 2017, respectively. In this case (where treatment is binary), the coefficient on $Post \times Treat$ (α_1) is our TWFE estimate for the intent-to-treat effect of CIM due to living in treated district.

---

[9] We use lagged coverage for 2016 and 2018 ASER rounds.

Specifically, α_1 is the conditional average test score difference between children in treatment and control districts in the post-treatment years.

Second, we use a continuous measure of treatment status, and define $Treat_{dt}$ as the number of toilets per 1000 of the district population constructed by the program in district d and year t. In regressions that utilize the continuous coverage of the program, treatment is represented by $Coverage\ per\ 1000$. In this case (where treatment is continuous), the intent-to-treat effect (α_1) retrieves the average linear effect of one additional CIM toilet in the district of residence, on children's math and literacy scores.

In equation (1), X_i is a matrix of child-level controls, that include the gender, age, age^2, and whether the child's parents attended school. Due to differences in the duration of CIM exposure and growth rates across children in different ages (and born in different years) there may be differences in the cognition changes from the program Yang (2011). Therefore, we include birth-year fixed effects, θ_b. In the equation, we also include θ_d and θ_t, representing district and time fixed effects. Lastly, $\theta_d \times t$ stands for district-specific time trends, accounting for the linear district-specific evolution of test-scores overtime.

4.2 Threats to identification

The primary assumption required for the TWFE model is that of parallel trends. Parallel trends means that without CIM, trends in children's math and literacy scores in treatment districts would have been identical to those in control districts. 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 outcome changes between the two groups cannot be entirely attributed to the program. We address this concern by checking for pre-treatment parallel trends using an event study. Failing this test indicates that our TWFE 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 our estimated intent-to-treat effect to this violation using the honest difference-in-differences approach developed in Rambachan and Roth (2023). We also utilize the continuous coverage of the program, and test if a quasi-increase of CIM coverage affects math and literacy scores prior to the implementation of the program in 2014.

An emerging area of literature highlights the potential challenges associated with estimating dynamic treatment effects through TWFE models, especially if an intervention is rolled-out in a staggered manner. Goodman-Bacon (2021) shows that treatment effects, in cases where a program is heterogeneously implemented across time (rolled-out in a staggered manner), can be biased away from a weighted average of the intent-to-treat effects on the treated. This occurs because past treated units are used as controls against the future treated units. These concerns are also highlighted by others (Borusyak et al., 2021; Roth, 2022; De Chaisemartin and d'Haultfoeuille, 2020; Sun and Abraham, 2021; Callaway and Sant'Anna, 2021; Athey et al., 2016), who explain that the canonical TWFE models fail to provide unbiased treatment effect estimates in absence of a uniform implementation of a given program.

To address bias due to related to staggered rollout some have proposed alternative estimators. For example, Callaway and Sant'Anna (2021), Wooldridge (2021), and Sun and Abraham (2021) propose estimating group-specific effects based on the timing of program uptake. Borusyak et al. (2021) obtain event study estimates by imputing untreated outcomes for treated observations, and then by calculating treatment effects as the weighted averages of the differences between actual and imputed outcomes. In this paper, to address the staggered nature of the implementation of CIM and to estimate the dynamic treatment effects of the program, we adopt the method proposed by Wooldridge (2021).

Another threat to identification in our TWFE strategy is the possibility of other programs being implemented at the same time as CIM that also can impact children's cognitive outcomes. The concern is that if there are other village-level programs affecting children's cognition, we may be falsely attributing the math and reading improvements to CIM. While we are unable to account for specific government programs, we examine this by estimating equation (1) on relevant village-level outcomes. These include whether the village has a primary health center, a bank, and a post office.

Similarly, to support the hypothesis that the cognition improvements are indeed driven by CIM, 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 regularly attends school, and whether household lives in a thatched house.

The validity of our research design also requires the exogeneity of the roll-out of CIM (i.e., there is not a non-random program placement). This means that the timing of CIM roll-out is not determined by household-/community-/village-level characteristics including pre-existing educational outcomes and toilet coverage. We address this concern by checking if pre-treatment characteristics (aggregated at the district-level) predict program roll-out. If we find that these characteristics significantly affect CIM implementation, then our TWFE results are likely biased due to non-random program placement.

Finally, we check to see if our conclusions are sensitive to how we define our test score variable. 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 and binary math and literacy scores as dependent variables. All of these checks largely support our main findings and are discussed in further detail in Section 5.

4.3 Results

Two way fixed effects estimates for math and literacy are presented in Table 2. In columns 1-4, our control group comprises of children in districts that are never exposed to CIM (i.e., by the end of our study, there was no CIM-constructed toilet in these districts). In columns 5 and 6, we include children in districts where the first CIM toilet was not constructed until 2018 within the control group. Therefore, columns 5 and 6 report the intent-to treat effects of CIM on children residing in early adoption districts (2014 and 2016) relative those in to late (2018)/no adoption districts.[10] In

---

[10] Early adopters sustain their first exposures in 2014, and 2015, and Late adopters sustain their first exposure in 2017.

columns 1 and 2, we exclude the control variables, which are included in all the other columns.

In columns 3 and 4, we find that living in a CIM-exposed district leads to positive and statistically significant improvements in children's standardized test-scores by approximately 0.033 standard deviation (SD) and 0.026 SD for standardized math and literacy scores, respectively. When we include late-adoption districts in the control group, we find almost identical intent-to-treat effect estimates in magnitude and statistical significance. Coefficients in columns 5 and 6 indicate that children in early-adoption CIM districts (2014, and 2016) report significant improvements in standardized math scores by 0.037 SD and literacy scores by 0.027 SD, than those in late adoption (2018) or non-adoption districts.

In Table 3, we use the number of toilets constructed through CIM in the district as our treatment variable.[11] This is indicated by $Coverage\ per\ 1000$ in the table. $Coverage\ per\ 1000$ represents the number of toilets constructed under CIM divided by the total district population, multiplied by 1000.[12] We find positive improvements on both math and literacy scores. Results indicate that the construction of 1 toilet per 1000 population in the district leads to significant improvements in children's math scores by 0.0012 SD. The coefficient for literacy is statistically insignificant and quite small in magnitude. Table 3 also reports that, on average, 12.92 household toilets are constructed for every 1000 rural population in the district, with a standard deviation of 16.59. Therefore, if we scale the effect by the average (standard deviation), our model predicts that constructing 12.92 (16.59) per 1000 CIM toilets will improve math scores by 0.017 (0.021) SD.[13]

We next examine whether there is heterogeneity in the benefits of CIM exposure across boys and girls, and different age groups, given evidence that cognitive production or human capital investment behavior can differ 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). These results are reported in Table 4. In columns 1 - 4, treatment status is defined as a binary indicator, and in columns 5 - 8, treatment

---

[11] In the continuous treatment models, we do not include late adoption (2018) districts within the control group.

[12] Total population in district is obtained from the 2010/11 Census.

[13] In Appendix Table A.5, we include the quadratic of Coverage per 1000 in the continuous coverage regressions.

is defined as the number of CIM-constructed toilets per 1000 population. In Panel A, we show the intent-to-treat estimates for boys and girls, and in Panel B, we show the results for two different age groups: 5 – 10 years and 11 – 16 years.

We do not find a substantial difference in the benefits of CIM to test performance across boys and girls. Estimated effects of binary treatment status (columns 1-2, Table 4) indicate that district-level exposure to CIM in early adoption districts causes significantly positive improvements in boys' and girls' math scores by 0.035 SD and 0.037 SD, relative to the late/non-adoption districts. In columns 5 and 6, an increase in $Coverage\ per\ 1000$ by 1 toilet leads to significant improvements in math scores by approximately 0.0015 and 0.0011 SD for both boys and girls, respectively (columns 5-6, Table 4). Intent-to-treat effect estimates for literacy in columns 3, 4, 7 and 8 are positive but very noisy. The only statistically significant estimate is that for the binary treatment effect on boys' literacy scores, with a point estimate of 0.033 SD. Since the average number of toilets constructed for the sample of boys and girls are 12.83 and 13.02 per 1000, respectively, our model predicts that math and literacy scores will improve by 0.019 SD, for boys (0.0014 multiplied with 12.83), and 0.013 SD for girls (0.001 multiplied with 13.02), on average across districts.

In panel B of Table 4, we examine the effect of CIM on test-scores separately for younger children (aged 5 – 10 years) and older children (aged 11 – 16 years). TWFE coefficients (columns 1 – 4) indicate that the intent-to-treat of residing in early adoption (2014 and 2016) districts are positive and statistically significant for both age cohorts. For both younger and older children in 2014 and 2016 districts, we find that CIM exposure results in improved math scores, by 0.046 SD and 0.029 SD, respectively, relative to those in the late adoption (2018)/non-adoption districts. The corresponding point estimates for literacy outcomes are 0.03 SD and 0.025 SD, respectively. In columns 5 and 6, we find that an increase in toilet construction per 1000 population in the district leads to significant and small improvements in math scores. These improvements are 0.0007 SD for children aged between 5 and 10 years, and 0.001 SD for those aged between 11 and 15 years. In columns 7 and 8, we also find that literacy scores improve significantly by 0.0003 SD for older children with an increase in toilet construction per 1000 population, but the estimated impact on literacy scores of younger children is not statistically significant. Based on the mean coverage estimates per 1000 population (12.89 for 5 – 10-year-olds and 12.97 for 11 – 16-year-olds), predicted average improvement in math is 0.0128 SD for those aged between 5 and 10 years, and 0.012 SD for those aged between 11 and 16 years.

Our results indicate that CIM causes significant improvements in children's math scores. Estimated impacts on literacy scores are positive, but those are not statistically significant for the continuous coverage models. We also show that children in early adoption districts report significantly improved test-scores than those in the late/non adoption districts. From the continuous coverage models, we find that an increase in toilet construction through the program causes significant improvements in children's math scores. Sub-group analysis estimates indicate that both boys and girls, and younger and older children experience positive and significant math improvements due to their exposure to CIM. This is indicated both in the two way fixed effects and continuous models.

Literacy scores do not improve significantly for both boys and girls, due to an increase in program coverage per 1000 population. But we do find significantly larger improvements in literacy scores for boys in the TWFE model. Results also indicate that an increase in the coverage of the program causes significant improvements in the literacy scores of older children. For those aged between 5 and 10 years, we only find a positive treatment effect in our TWFE model. This coefficient is statistically significant at the 10 percent level.[14]

4.4 Event study

Existing research like Miller et al. (2021), Clarke and Tapia-Schythe (2021), Cunningham (2021), Freyaldenhoven et al. (2019) recommend using event study designs to evaluate treatment effects, for two different reasons. One, they allow us to check for parallel pre-treatment trends between the treatment and control groups, failing which, we cannot causally link the post-treatment outcome changes and the event (CIM in our case). Two, they help us examine the time of appearance of

---

[14] Appendix Tables A.3, A.4, A.6, and A.7 present main results using the balanced sample of districts.

treatment effects, whether the effects are increasing or decreasing overtime, and whether the effects are permanent or transient. Following these arguments, equation (2) is our event study design:

$$Z_{ibdt} = \delta_0 + \sum_{j=-7}^{-1} \delta_j.(Leads\ j)_{dt} + \sum_{k=0}^{2} \delta_k.(Lags\ k)_{dt} + \delta_3.X_i + \theta_d + \theta_t + \theta_b + \theta_d \times t + \omega_{ibdt} \tag{2}$$

On the right of equation 2, $Leads$ and $Lags$ are binary indicators of the number of period (year) differences between the year of implementation of CIM in the district, and the current year. We include 7 $Leads$ (for each period leading to the implementation of CIM) and 2 $Lags$ (for periods following the implementation of CIM) in the model, with the lead for period -1 being the comparison (base) year. The total set of $Leads$ and $Lags$ range from -8 to +2, however, we trim the number of pre-treatment periods by combining leads -7 and -8 together. Therefore, from equation (2), we obtain estimates for 7 periods before CIM, and 3 periods following CIM. Each coefficient on $\delta_{j\ or\ k}$ present the effect of potential exposure to CIM for children in treated districts relative to the year immediately prior to the year of implementation of the program.

Event study estimates for math and literacy are presented in Table 5 and plotted in Figure 4.[15] Vertical axes in panels A and B of the figure show the intent-to-treat effects of the treated districts in each period, relative to the year immediately before CIM was implemented (represented by the dashed red vertical line). Y axes characterize the pre ($Leads$) and post ($Lags$) treatment periods. For math, the pre-treatment coefficients are negative, but they are statistically significant. This tells us that prior to the implementation of CIM, children in treated districts report significantly different trends in math scores, relative to controlled districts. But for the post-treatment periods, we find positive and significant differences in math scores by around 0.748 SD, and 1.444 SD, between the treatment and control districts. These coefficients are for Lag periods +2 and +4. The coefficient for period zero is positive (0.204 SD), but not statistically significant.

Since the $Lead$ (pre-treatment) estimates are statistically significant, we cannot rule out pre-existing differential trends in math scores between treatment and control districts. However, these pre-treatment estimates are negative, which demonstrates that before CIM, children in treated dis-

---

[15] Results for balanced sample are shown in Table A.6 and plotted in Figure A.1.

tricts report significantly smaller differences in math scores than those in control districts. With the implementation of CIM, these differences are positive. We also find that the positive intent-to-treat effects of CIM on math scores become larger and gain statistical significance as the duration of exposure to the program increases every two years.

Estimates for literacy, in column 2 of Table 5 and panel B of Figure 4 are statistically insignificant in both the pre- and post-treatment periods. Similar to the estimates for math, all the $Lead$ (pre-treatment) coefficients are negative, and the Lag (post-treatment) coefficients are positive and increasing, although not statistically significant. Statistically insignificant coefficients for the $Lead$ periods show that prior to the implementation of CIM, children in treatment districts report similar changes (trends) in literacy scores than those in the control districts.

5 Robustness

5.1 Violation of parallel pre-treatment trends

Post-treatment differential trends stand for the differences between children in treatment and control districts, in their expected time-differences in math and literacy scores. These time-differences are between the base period (-1) and the post-treatment periods (+2, or 0, or +4). Because post-treatment differential trends are counterfactual, i.e., intended to be calculated in absence of CIM, pre-treatment differential trends are informative about post-treatment differential trends. Based on this assumption, we use a standard event study design (in section 4.4) to test if pre-treatment differential trends are statistically similar to zero, which tells us whether parallel trends are valid?[16]

Parallel trends are violated in our event study analysis. This means that the two way fixed effects (TWFE) results are potentially biased. However, failing to reject parallel trends, as shown in Figure A.1 and Table A.7, may not be similar to confirming it (Kahn-Lang and Lang, 2020; Marcus and Sant'Anna, 2021). To address this violation of parallel pre-treatment trends, we adopt

---

[16] Pre-treatment differential trend stands for differences between treatment and control districts in average outcomes, between the pre-treatment period and base period (-1).

the relative magnitudes approach developed by Rambachan and Roth (2023). Using this method, we test for the sensitivity of the intent-to-treat effect estimate for period +2 in our event study findings (Table 5 and Figure 4).

The relative magnitudes method relates to the understanding that pre-existing district specific programs/shocks (e.g., government programs, or school-level changes) affect cognition outcomes of children in treated districts differently than those in control districts. Because these programs are pre-existing, their cognition effects can differ across the post-treatment (0, +2, and +4) and pre-treatment periods (-7, -6, -5, -4, -3, and -2).[17] To account for these differences, following Rambachan and Roth (2023), we introduce a relative magnitudes parameter ($M \leq 0$).[18] This parameter stands for the proportion of the maximal (or largest) pre-treatment differential trends, that post-treatment differential trends (which is counterfactual) can be equivalent to.

This is formalized as: $Post_{trends} = M \times Pre_{trends}$. $Post_{trends}$ represents post-treatment differential trends, and Pre_{trends} represents the maximal (or largest) pre-treatment differential trends. M equals 1 if differential trends are assumed to be constant across pre- and post-treatment periods (i.e., $Post_{trends} = Pre_{trends}$). Other values of M that we use are 0.5, 1.5, and 2. M=1.5 or 2 indicate that $Post_{trends}$ are 1.5 or 2 times the largest 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 these values, we obtain a breakdown point of M. At this breakdown point, the CIM treatment effect for period +2 is zero (i.e., no longer statistically significant). This means that at the breakdown point, the differential cognition effect of district specific pre-existing programs, are large enough for the effect of CIM in period +2 to be statistically similar to zero.

Using this strategy, we obtain multiple confidence intervals for the intent-to-treat estimates for the post-treatment period, +2, from our event study. These intent-to-treat estimates (from Table

---

[17] This means that non-parallel (differential) trends in children's cognition, can differ across the post- and pre-treatment periods.

[18] Relative magnitude is the ratio of the post-treatment differential trend (unobserved) and the largest pre-treatment differential trend (pre-treatment coefficient that is statistically significant). Statistical significance implies a violation of parallel trends.

5 and Figure 4) are 0.748 SD and 0.268 SD for math and literacy, respectively. The confidence intervals tell us if the CIM treatment effects for period +2 are sensitive to relative changes in pre-treatment differential trends, between the pre- and post-treatment periods. Therefore, using the relative magnitudes method we measure relative changes in the differential effect of pre-existing programs, for us to reject a null effect (of CIM). These relative changes in the differential effect, between pre-treatment and post-treatment periods, can occur due to different cultural, political, economic, and environmental factors (e.g., elections, government policies, natural or health emergencies).

The confidence intervals are shown in Figure 5 and reported in Table A.8. For math scores, shown in Panel A of the figure, confidence intervals (0.172 and 1.323) from the event study results (Table 5) for period +2 are shown in red. Blue bars show the confidence intervals of the intent-to-treat effects for period +2 for different values of M. We find that if post-treatment differential trends are restricted to equaling the maximal pre-treatment differential trends (i.e., M=1), the intent-to-treat estimate for period +2 is positive and statistically significant. The lower and upper bounds for this coefficient are 0.069 and 1.584, respectively. Looking further to the right, the breakdown point of M is 1.5 [CI= -0.007, + 1.764]. This indicates that post-treatment differential trends need to be equivalent to 1.5 times as large as the maximal pre-treatment differential trends, for the treatment effect of CIM to be statistically insignificant. This tells us that for the effect of CIM to be zero, the effect of pre-treatment programs in post-treatment period (+2), need to be almost twice as much as their largest effect in the pre-treatment period.

To put in perspective, from our event study design we find that the largest statistically significant pre-treatment trend estimate is 0.67 SD (absolute value). Therefore, for a null treatment effect of CIM in period +2, the differential cognition effect of pre-existing programs must be approximately 1.005 SD. The post-treatment period +2 stands for the years 2016 or 2018.[19] Given that multiple state-level elections were held during these years (2016 and 2018), it is plausible that differences in math scores, due to pre-existing programs, between children in treatment and control

---

[19] This is because 519 districts first sustain exposure to CIM in 2014, and 58 first sustain exposure to the program in 2015. We use the coverage information from previous year.

districts potentially widened due to these elections or policies implemented by the elected parties.[20] This means that the effect of these elections can underly the positive intent-to-treat effects we estimate in our event study. However, evidence showing that the effect of pre-existing programs on children's cognition increased (multiplied by 1.5 points) due to these elections is limited. This tells that pre-treatment differential trends in math scores are likely to be similar across the pre- and post-treatment periods. Since the respective confidence intervals for M=1 are positive (statistically significant), pre-treatment differential trends, or pre-existing cognition differences treated and controlled children, are unlikely to bias the intent-to-treat effects we estimate in the TWFE, continuous coverage, and event study strategies. Estimated confidence intervals for literacy, shown in Panel B of Figure 5, are statistically insignificant for all values of M.[21] Therefore, children's literacy scores do not improve significantly due to CIM, even if we account for the effect of pre-treatment district specific programs on children's cognition.

5.2 Effect of program coverage in pre-treatment period

We adopt another strategy using the continuous coverage of the program from the first year and from the last (most recent) year of treatment. In this strategy, we just use the pre-treatment ASER surveys (2010, 2011, 2012, and 2013), and estimate the quasi effects of toilet construction per 1000 population on children's math and literacy outcomes, in absence of actual treatment. This is presented in equation (3):

$$Z_{ibdt} = \beta_0 + \beta_1 . Coverage\ per\ 1000_{dt} + \sum_{p=-6}^{-1} \beta_{pdt} . (Leads\ p \times Coverage\ per\ 1000)_{dt} + \beta_3 . X_i + \theta_d + \theta_t + \theta_b + \theta_d \times t + \upsilon_{ibdt} \tag{3}$$

where, similar to equation (2), $Leads\ p$ are binary indicators of the number of period (year) differences between the first year of implementation (coverage) of CIM in the district and the

---

[20] Elections in 2018 – Meghalaya, Nagaland, Mizoram, Madhya Pradesh, Chhattisgarh, Rajasthan, Karnataka, and Telangana. Elections in 2016 – West Bengal, Assam, Puducherry, Kerala, and Tamil Nadu.

[21] Results based on a different approach, the smoothness bounds strategy, are available in the appendix.

current year. We focus on the pre-treatment ASER rounds (2010, 2011, 2012, and 2013), so we only include $Lead$ periods in this equation. These $Lead$ periods range from -8 to -1, where $Leads$ -8 and -7 are combined together. In this equation, $Coverage\ per\ 1000$ represents the total number of toilets constructed in the district per 1000 population, in the first year of coverage or in the most recent (last) year of coverage. Therefore, coefficients on the interactions between $Leads\ p$ and $Coverage\ per\ 1000$ (β_{pdt}), indicate the quasi effect of toilet construction, prior to implementation of the program, in each of the pre-treatment periods. These coefficients tell us the quasi effect of the program on math and literacy, for children in districts with high coverage relative to those with low coverage of the program. β_1 is the quasi effect of toilet construction in the earliest pre-treatment period (-7 and -8 leads).

It is worth noting that equation (3) does not estimate cognition differences between the treatment and control groups. Rather we are estimating the quasi effects of program coverage in each of the pre-treatment periods, i.e., before the actual implementation of the program. This means that β_1 represents the effect of $Coverage\ per\ 1000$ in the earliest pre-treatment period (-7 and -8 combined), and coefficients on $Leads\ p \times Coverage\ per\ 1000$ show the effect of $Coverage\ per\ 1000$ in periods -6, -5, -4, -3, -2, and -1. Our strategy is motivated by Basu (2021), Muralidharan and Prakash (2017), and Suryanarayana et al. (2023). However, we adopt a slightly different approach from theirs. This is because, instead of using a continuous pre-treatment year variable, which is interacted with the coverage data, we use binary indicators for each pre-treatment period. This method gives the advantage of estimating the effect of program coverage in each pre-treatment period.

Results are shown in Table 6 and plotted in Figures 6, using the first year's coverage, and 7, using the most recent year's coverage.[22] Vertical axes in panels A and B of the figures show the quasi effects of toilet construction per 1000 population in the district. Horizontal axes characterize the pre ($Leads$) treatment periods for each district. Estimated coefficients on $Coverage\ per\ 1000$ are available in the table but those are not plotted in the figures. In Figure 6, where we interact

---

[22] Coefficients on $Coverage\ per\ 1000$ are reported in Table 6.

each pre-treatment period with program coverage from the first treatment year, the pre-treatment coefficient is positive and statistically significant in period -6, but all the remaining coefficients are negative. Estimated effects for periods -4, and -5 are statistically insignificant, but those for periods -3, -2, and -1 are statistically significant. These results indicate that the quasi effect of toilet construction on children's math scores are mostly negative. These negative effects are statistically insignificant for two periods, and they decline further as the duration towards the initial implementation of CIM reduces (i.e., in periods -3, -2, and -1). Literacy estimates, for periods -6, -5, -4, -3, -2, and -1, in Panel B, are positive but not statistically significant. This tells us that, in absence of actual treatment, the (false) construction of toilets does not cause significant improvements in literacy scores.

Estimated coefficients in Figure 6, where we interact each pre-treatment period with program coverage from the most recent year, represent a similar scenario. All the coefficients are statistically insignificant for both math and literacy scores, except for the literacy coefficient in period -6. In both Panels A and B in the figure, all the coefficients are close to zero. Math estimates are mostly positive, whereas those for literacy are negative, although statistically insignificant.

To summarize, the estimated impacts of CIM on children's math and literacy scores, in absence of actual treatment (i.e., in the pre-treatment period), are mostly statistically insignificant. These results are shown in Figures 6 and 7. This is the opposite to the positive and significant intent-to-treat effects we find in our TWFE and continuous coverage regressions. In Panel A of Figure 6, we also find that the quasi effects of toilet construction on math scores are negative. This negative impact becomes larger as we move closer to the first year of implementation of the program.

5.3 Dynamic treatment effects

If different units (districts) are treated in different periods, treatment effects obtained from a two way fixed effects framework are biased and they may fail to identify the causal effect of the treatment. This happens because early treatment districts are used as controls for those that receive treatment in later periods. To address this concern, we adopt the method developed in Wooldridge

(2021). This method allows us to exploit the staggered nature of implementation of the program and recover dynamic improvements in children's math and literacy scores. Equation (4) presents the strategy:

$$Z_{ibdt} = \sigma_0 + \sum_{g=2014}^{2016} \sum_{\tau=2014}^{2018} \sigma_{g\tau} \times 1(g,\tau) + \theta_d + \theta_t + \theta_b + \theta_d \times t + \zeta_{ibdt} \tag{4}$$

where, g is the first year of exposure to CIM, which is either 2014 or 2016, τ represents each of the post-treatment years of 2014, 2016 or 2018. $\sigma_{g\tau}$ corresponds to each interaction between g and τ, and the treatment effect indicator $1(g,\tau)$, that takes a value of 1 if a district is ever treated in 2014 or 2016. As an alternative to the binary treatment effect indicator, we also use the continuous coverage of the program, i.e., the number of toilets constructed per 1000 population in the district. In both these regressions, $\sigma_{g=2014,\ \tau=2014,\ 2016,\ 2018}$ give us the dynamic math and literacy improvements in 2014, 2016 and 2018, for districts that are first exposed to CIM in 2014, as compared to the districts that did not yet sustain exposure in 2014. Similarly, $\sigma_{g=2016,\ \tau=2016,\ 2018}$ estimate the dynamic effects in 2016 and 2018, for districts that adopted the program in 2016, relative to the never treated districts (and the not yet treated 2018 districts). In sum, $\sigma_{g\tau}$ represent the intent-to-treat effect of districts first treated in 2014, in each of the post-treatment years (2014, 2016, and 2018); or the intent-to-treat effect of districts first treated in 2016, in each of the post-treatment years (2016, and 2018). In this equation, we include district, year, and birth-year fixed effects, and district-year trends, and exclude the control variables.[23]

Results are shown in Table 7. Figure 8 plots the dynamic intent-to-treat effects for districts that adopted the program in 2014. In the table, columns 1 and 2 present the estimates for the binary treatment effect models, and columns 3 and 4 are for the continuous treatment models. For math, in column 1, districts that adopted CIM in 2014, report positive and statistically significant effect in 2016 and 2018 by 0.093 SD and 0.182 SD. In the continuous treatment effect model (column 3), the intent-to-treat effect estimates in 2016 and 2018 are 0.001 SD and 0.002 SD respectively. Literacy estimates in column 2 also show that districts that adopted CIM in 2014, report positive

---

[23] Results with controls are shown in Appendix Table A.9

and statistically significant treatment estimates in 2016 and 2018 by 0.093 SD and 0.13 SD. But in the continuous treatment effect model (column 4), the coefficient in 2018 (0.005 SD) is slightly smaller than that for 2016. Coefficients for children in districts that adopted the program in 2016 are statistically insignificant and report large standard errors.

In panel A of Figure 8, we show that all the math and literacy (reading) intent-to-treat effect estimates of children in districts that implemented the program in 2014, are positive and statistically significant. These coefficients become larger with every 2 years of additional exposure to CIM. In panel B, we plot the continuous treatment effects coefficients. The impacts of CIM on math scores are positive and statistically significant for each year, except for 2014. This impact increases with an increase in children's exposure to the program. Coefficients for literacy scores do not show an increasing nature of impact of the program. This is represented by the smaller estimate in 2018 relative to 2016. Estimated coefficients for literacy are also statistically insignificant for 2016 and 2014. These results show that CIM causes significant and dynamic improvements in children's math scores, suggesting that a greater duration of exposure to improved sanitation facilities due to the program causes both instantaneous and dynamic changes in children's math scores.

Table 8 presents a comparison of our results. Intent-to-treat effects obtained from Wooldridge (2021) are shown in panel B. Those obtained from the two way fixed effects and the continuous treatment models, from Tables 2 and 3, are shown in panel A. Coefficients for math scores are positive and statistically significant across all the models. Estimates in column 1 are 0.0377 SD and 0.0379 SD. From the continuous coverage models, these estimates are 0.001 SD and 0.0004 SD. Coefficients for literacy scores are positive and statistically significant in column 2. These are not significant in column 4. Therefore, we find that treatment effects estimated in our main results (Tables 2 and 3), are closely identical to those calculated in Wooldridge (2021). These consistent findings indicate that CIM causes significant improvements in children's math scores. Coefficients for literacy scores are positive, but they do not consistently exhibit statistical significance.

5.4 Potential mechanisms

Through the Clean India Mission, improved household-level access to better sanitation facilities can reduce the likelihood of children being affected by short-term illnesses. This affects their school attendance positively (Battiston et al., 2013; Dreibelbis et al., 2013a,b; Rauniyar et al., 2011), which can lead to improvements in their cognition outcomes. The program is also associated with benefits like improved community and household cleanliness and living conditions, better household infrastructure and monetary value, and reduced exposure to pollution, (Hutton, 2013; Banerjee et al., 2017; Behera et al., 2021). All these factors can contribute to better cognition outcomes of children. Therefore, to test for potential mechanisms through which CIM influences math and literacy scores, we check whether the program leads to changes in household-level access to a toilet, improves school attendance of children, and reduces the likelihood of household living in a thatched house.[24] It is, however, noteworthy that none of these changes are direct mandates of CIM, but they act as potential channels through which the program can affect children's cognition outcomes.

In columns 1 – 3 of Table 9, we use the TWFE model to evaluate the effect of CIM on each of these outcomes. The probability of the household having its own toilet has improved significantly for treated children, relative to those in the control group. Coefficients for school attendance (School Att.) and living in a thatched house are not statistically significant. But in the continuous coverage models (columns 4 – 6), we find that household toilet construction per 1000 rural population yields a significant but small improvement in children's school attendance, and a significant decline in the probability of living in a thatched house. The negative estimates for thatched housing (except in column 3), and positive estimates for toilet access and school attendance, suggest that district level coverage of CIM causes positive changes in children's cognition through these factors.

---

[24] Thatched houses are associated with lack of access to proper sanitation facilities, lower socioeconomic statuses, and increased incidences of illnesses (Azurin and Alvero, 1974; Prithviraj et al., 2017)

5.5 Co-existing programs

The possibility that the government implemented other village-level programs that focus on improving people's access to government services is a concern in our study. This is because, both our TWFE and continuous coverage models could be picking up the positive effect of these improved access to government services on children's cognition, instead of CIM. To check this, we examine if the district level coverage of CIM predicts village-level access to a primary health center, and to organizations like a bank and post office.

These results are shown in Table 10. In columns 1 – 3 we run the TWFE model, and in columns 4 – 6 treatment is indicated by $Coverage\ per\ 1000$. The estimated treatment effects from this exercise are statistically insignificant and small in magnitude. It, therefore, seems unlikely that complementary government programs that improve village-level access to a primary health center, bank, or post office, underly the positive treatment effects we estimate for CIM.

5.6 Determinants of program implementation

Another concern with evaluating government programs relates to the endogenous nature of the program roll-out. This means that the implementation of CIM could be driven by different household, village, district-level factors. We address this concern examining the association between different pre-treatment household, village and district-level characteristics, and the number of toilets constructed (per 1000 population) in the district through CIM in its first year of implementation.

We use the pre-treatment ASER rounds (2010, 2011, 2012, and 2013), and aggregate all the household and village-level data to the district level. These data provide information on the percentage of villages in the district with a post-office, bank, and health center. At the household level, these variables are the percentage of households in district living in a thatched house, and percentage with toilet in the house. We also include data on percentage of school attending children, percentage of male children, mean age of children, and the percentage of parents who have attended formal schooling.

Before the implementation CIM in 2014, the Ministry of Drinking Water and Sanitation conducted a survey in 2013 in each district. This survey provides information on the total number of rural households in the district, total number of rural households with/without a toilet, total number of rural households above the poverty line (APL) without a toilet, and the total number of rural households below the poverty line (BPL) without a toilet. Using district-level populations, we create three different variables and include those in the regressions. These are the number of rural households without toilets per 1000 population, number of rural APL households without toilets per 1000 population, and the number of rural BPL households without toilets per 1000 population. Our strategy is similar to Bronchetti et al. (2019), Hoynes et al. (2011), Hoynes and Schanzenbach (2009), or Amuedo-Dorantes and Bucheli (2022).

Results are shown in Table 11. We find that none of the estimated coefficients (except the weakly significant coefficient on mean age of children) are statistically significant. Notably, there is no significant association between the percentage of households with an already existing toilet in the district and CIM coverage. This indicates that the district-level coverage of CIM is unlikely to be endogenous in nature, and our results are unlikely to represent biased treatment effects due to non-random program implementation.

5.7 Test score measurement

Standardized test scores are an unbiased method of measuring students' cognitive skills as they represent a comparison between the sample mean and individual test scores (Angrist and Lavy, 1997; Hunter and Hamilton, 2002). In the ASER surveys, however, math and literacy assessments represent children's learning proficiency rates, instead of test scores. These proficiency rates are 0, 1, 2, 3, and 4. However, the categorical measurements of these outcomes may not differ from the sample mean by large amounts. To address this concern, we use the raw and binary math and literacy scores as the dependent variables.

These results are shown in Table 12. In Panel A we use the raw math and literacy outcomes.[25]

---

[25] 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.

In Panel B, math equals 1 if the child can do three by one digit division (has level 4 proficiency) and it equals 0 for the other proficiency levels. Literacy (reading) equals 1 if the child has a level 4 proficiency, i.e., can read grade 2 text, and 0 for the other levels. In the TWFE regressions, shown in columns 1 and 2, the intent-to-treat effect estimates for both math and literacy are positive and statistically significant. For the continuous coverage models, in columns 3 and 4, the estimated effect is positive for both math and literacy scores, however they are only significant for math. These findings conform to our main result, that the construction of toilets through the Clean India Mission program fosters improved math learning achievements of children.

6 Conclusion

We examine the effect of a large-scale sanitation program on children's cognitive abilities in India. We use the conventional TWFE, continuous coverage, and event study methodologies to find significant instantaneous and dynamic improvements in math test-scores. Our findings are supported by different sub-group analysis, where we estimate similar improvements for both boys and girls, and for younger and older children. We also show that children in early adoption districts report significantly improved test-scores than those in the late/non adoption districts.

As an alternative to our event study findings, we adopt the honest difference-in-differences method. We also show that relative to the earliest pre-treatment period, the quasi-treatment effects of toilet construction on math scores are negative in absence of actual treatment. These negative impacts become larger as we move closer to the first year of implementation of the program. Therefore, pre-existing math trends do not pose a serious threat to the validity of our main results. We show that our results are not driven by co-existing government programs that can affect child cognition. The program has also improved children's school attendance, the probability of households living in a thatched house, and households' access to a toilet. These changes act as pathways through which CIM affects children's math and literacy scores.

By estimating positive changes in children's math scores due to the implementation 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 unintended 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 religion of children, which would allow us to examine the group specific effects of the program. We are also unable to match the ASER and CIM 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 math scores. This could potentially result in better human capital outcomes for these children in the future.

References

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.

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.

Amuedo-Dorantes, C. and Bucheli, J. R. (2022). Immigration policy and hispanic representation in national elections. *Journal of Policy Analysis and Management*.

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).

Angrist, J. D. and Lavy, V. (1997). The effect of a change in language of instruction on the returns to schooling in morocco. *Journal of Labor Economics*, 15(1, Part 2):S48–S76.

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.

Bartram, J., Lewis, K., Lenton, R., and Wright, A. (2005). Focusing on improved water and sanitation for health. *The Lancet*, 365(9461):810–812.

Basu, D. (2021). Majoritarian politics and hate crimes against religious minorities: Evidence from india, 2009–2018. *World Development*, 146:105540.

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.

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, 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.

Bronchetti, E. T., Christensen, G., and Hoynes, H. W. (2019). Local food prices, snap purchasing power, and child health. *Journal of health economics*, 68:102231.

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).

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.

Datt, G. and Ravallion, M. (2002). Is india's economic growth leaving the poor behind? *Journal of economic perspectives*, 16(3):89–108.

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.

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.

Hammer, J. and Spears, D. (2016). Village sanitation and child health: Effects and external validity in a randomized field experiment in rural india. *Journal of Health Economics*, 48:135–148.

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

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.

Hoynes, H., Page, M., and Stevens, A. H. (2011). Can targeted transfers improve birth outcomes?: Evidence from the introduction of the wic program. *Journal of Public Economics*, 95(7-8):813–827.

Hoynes, H. W. and Schanzenbach, D. W. (2009). Consumption responses to in-kind transfers: Evidence from the introduction of the food stamp program. *American Economic Journal: Applied Economics*, 1(4):109–139.

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

Hunter, J. E. and Hamilton, M. A. (2002). The advantages of using standardized scores in causal analysis. *Human Communication Research*, 28(4):552–561.

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.

Jasper, C., Le, T.-T., and Bartram, J. (2012). Water and sanitation in schools: a systematic review of the health and educational outcomes. *International journal of environmental research and public health*, 9(8):2772–2787.

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.

Kahn-Lang, A. and Lang, K. (2020). The promise and pitfalls of differences-in-differences: Reflections on 16 and pregnant and other applications. *Journal of Business & Economic Statistics*, 38(3):613–620.

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.

Marcus, M. and Sant'Anna, P. H. (2021). The role of parallel trends in event study settings: An application to environmental economics. *Journal of the Association of Environmental and Resource Economists*, 8(2):235–275.

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

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.

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*, page rdad018.

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.

Ravallion, M. and Datt, G. (2002). Why has economic growth been more pro-poor in some states of india than others? *Journal of development economics*, 68(2):381–400.

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.

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.

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.

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.

Suryadarma, D., Suryahadi, A., Sumarto, S., and Rogers, F. H. (2006). Improving student performance in public primary schools in developing countries: Evidence from indonesia. *Education Economics*, 14(4):401–429.

Suryanarayana, R., Chaudhuri, K., Nilayamgode, M., and Jha, N. (2023). Alcohol ban and crime: The abc's of the bihar prohibition.

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*.

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*.

Yang, Y. (2011). Aging, cohorts, and methods. In *Handbook of aging and the social sciences*, pages 17–30. Elsevier.

Table 1: Average characteristics across ASER and CIM rounds

	2010 N=402,292 (1)	2011 N=380,890 (2)	2012 N=344,230 (3)	2013 N=340,654 (4)	2014 N=336,584 (5)	2016 N=328,279 (6)	2018 N=323,858 (7)
Math	0.189	0.069	-0.050	-0.066	-0.078	-0.071	-0.042
	(1.002)	(1.001)	(0.992)	(0.998)	(0.999)	(0.991)	(0.982)
Literacy	0.112	0.047	-0.040	-0.045	-0.050	-0.044	-0.006
	(0.933)	(0.964)	(1.006)	(1.019)	(1.035)	(1.029)	(1.016)
Age in years	10.378	10.409	10.328	10.288	10.291	10.292	10.224
	(3.236)	(3.258)	(3.261)	(3.271)	(3.263)	(3.284)	(3.265)
Male (1=Yes, 0=No)	0.545	0.529	0.516	0.513	0.509	0.501	0.500
	(0.498)	(0.499)	(0.500)	(0.500)	(0.500)	(0.500)	(0.500)
Father attended school (1=Yes, 0=No)	0.723	0.730	0.735	0.734	0.741	0.743	0.764
	(0.447)	(0.444)	(0.441)	(0.442)	(0.438)	(0.437)	(0.425)
Mother attended school (1=Yes, 0=No)	0.518	0.509	0.505	0.518	0.530	0.559	0.594
	(0.500)	(0.500)	(0.500)	(0.500)	(0.499)	(0.497)	(0.491)
Thatched house (1=Yes, 0=No)	0.355	0.350	0.341	0.333	0.318	0.309	0.266
	(0.478)	(0.477)	(0.474)	(0.471)	(0.466)	(0.462)	(0.442)
Toilet in house (1=Yes, 0=No)	0.416	0.412	0.404	0.424	0.458	0.546	0.726
	(0.493)	(0.492)	(0.491)	(0.494)	(0.498)	(0.498)	(0.446)
Attends school (1=Yes, 0=No)	0.963	0.964	0.968	0.972	0.972	0.971	0.975
	(0.188)	(0.187)	(0.175)	(0.165)	(0.165)	(0.167)	(0.157)
Health center in village (1=Yes, 0=No)	0.419	0.434	0.442	0.429	0.425	0.423	0.412
	(0.493)	(0.496)	(0.497)	(0.495)	(0.494)	(0.494)	(0.492)
Bank in village (1=Yes, 0=No)	0.225	0.236	0.252	0.252	0.262	0.291	0.294
	(0.418)	(0.425)	(0.434)	(0.434)	(0.440)	(0.454)	(0.455)
Post office in village (1=Yes, 0=No)	0.449	0.443	0.448	0.419	0.417	0.405	0.382
	(0.497)	(0.497)	(0.497)	(0.493)	(0.493)	(0.491)	(0.486)
CIM coverage in district in current year					8,438.385	19,232.262	48,984.909
					(10,076.228)	(22,802.816)	(47,553.723)
CIM coverage in district in previous year						9,082.455	32,745.885
						(13,669.087)	(37,519.630)
CIM coverage per 1000 in current year					6.731	15.472	34.820
					(8.499)	(14.150)	(30.322)
CIM coverage per 1000 in previous year						6.974	25.392
						(7.731)	(21.765)
Percent rural households with own toilet				35.545	38.988	49.347	80.303
				(26.257)	(27.094)	(27.893)	(35.026)
Percent rural population with own toilet				7.053	7.694	9.952	15.903
				(5.456)	(5.576)	(5.875)	(6.539)

Notes: CIM coverage means the number of toilets constructed under the program in the district. CIM coverage in previous year for 2018, and 2016 are for 2017, and 2015. Data obtained from: Website

Table 2: Two way fixed effects

					Early v. Late/No	
	(1) Math	(2) Literacy	(3) Math	(4) Literacy	(5) Math	(6) Literacy
TWFE	0.0309^{*}	0.0239	0.0336^{***}	0.0260^{**}	0.0377^{***}	0.0273^{**}
	(0.0184)	(0.0198)	(0.0115)	(0.0117)	(0.0119)	(0.0118)
Constant	32.67	-4.227	40.30	4.875	40.32	4.883
	(38.35)	(39.18)	(25.79)	(20.40)	(25.79)	(20.40)
N	2456787	2456787	2456787	2456787	2456787	2456787
R^2	0.393	0.415	0.448	0.467	0.448	0.467
Covariates			Yes	Yes	Yes	Yes

Notes: All models include district fixed effects, year fixed effects, district-year trends, and birth-year fixed effects. Covariates - parental education, age, age^2, and gender. Standard errors clustered by district and birth year in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 3: Continuous treatment effects

	(1) Math	(2) Literacy
Coverage per 1000	0.00128^{***}	0.000115
	(0.000171)	(0.000169)
Constant	33.56	4.233
	(25.86)	(20.44)
N	2456787	2456787
R^2	0.449	0.467
Mean coverage per 1000	12.92	12.92
Standard deviation	16.59	16.59

Notes: Both models include district fixed effects, year fixed effects, district-year trends, and birth-year fixed effects. Covariates - parental education, age, age^2, and gender. Standard errors clustered by district and birth year in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 4: Heterogeneity across gender and age groups

	Math		Literacy		Math		Literacy	
Panel A: Gender	(1) Boys	(2) Girls	(3) Boys	(4) Girls	(5) Boys	(6) Girls	(7) Boys	(8) Girls
TWFE	0.0358***	0.0378**	0.0334**	0.0189				
	(0.0127)	(0.0149)	(0.0137)	(0.0142)				
Coverage per 1000					0.00149***	0.00109***	0.000238	0.00000426
					(0.000210)	(0.000210)	(0.000212)	(0.000201)
Constant	24.71	52.02	7.224	-1.675	17.12	46.15	5.948	-1.732
	(22.89)	(36.47)	(21.40)	(28.63)	(23.01)	(36.52)	(21.45)	(28.67)
N	1270745	1186042	1270745	1186042	1270745	1186042	1270745	1186042
R^2	0.465	0.436	0.474	0.463	0.465	0.436	0.474	0.463
Mean coverage per 1000					12.83	13.02	12.83	13.02
Standard deviation					16.50	16.69	16.50	16.69

Panel B: Age groups	(1) 5 - 10 years	(2) 11 - 16 years	(3) 5 - 10 years	(4) 11 - 16 years	(5) 5 - 10 years	(6) 11 - 16 years	(7) 5 - 10 years	(8) 11 - 16 years
TWFE	0.0464***	0.0297*	0.0302*	0.0252**				
	(0.0151)	(0.0156)	(0.0183)	(0.0115)				
Coverage per 1000					0.000741***	0.00184***	-0.0000880	0.000372*
					(0.000225)	(0.000228)	(0.000250)	(0.000193)
Constant	29.17	54.32**	-9.150	21.88	25.21	44.78**	-8.739	19.91
	(35.40)	(22.04)	(29.86)	(15.31)	(35.56)	(22.22)	(29.89)	(15.35)
N	1301016	1155771	1301016	1155771	1301016	1155771	1301016	1155771
R^2	0.348	0.111	0.349	0.0823	0.348	0.111	0.349	0.0823
Mean coverage per 1000					12.89	12.97	12.89	12.97
Standard deviation					16.65	16.53	16.65	16.53

Notes: All models include district fixed effects, year fixed effects, district-year trends, and birth-year fixed effects. Covariates - parental education, age, age^2, and gender. Standard errors clustered by district and birth year in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 5: Event study

	(1) Math	(2) Literacy
-7	-0.786*	-0.186
	(0.432)	(0.587)
-6	-0.672**	-0.284
	(0.329)	(0.451)
-5	-0.663**	-0.175
	(0.296)	(0.417)
-4	-0.476**	-0.238
	(0.201)	(0.270)
-3	-0.372**	-0.119
	(0.156)	(0.222)
-2	-0.236***	-0.146*
	(0.0714)	(0.0747)
0	0.204*	0.0343
	(0.118)	(0.178)
+2	0.748**	0.268
	(0.294)	(0.438)
+4	1.444***	0.611
	(0.500)	(0.730)
Constant	40.75	5.277
	(25.79)	(20.40)
N	2456787	2456787
R^2	0.449	0.467

Notes: Both models include district fixed effects, year fixed effects, district-year trends, and birth-year fixed effects. Covariates - parental education, age, age^2, and gender. Standard errors clustered by district and birth year in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 6: Effect of coverage on pre-treatment outcomes

	First year's coverage		Most recent year's coverage	
	(1) Math	(2) Literacy	(3) Math	(4) Literacy
Coverage per 1000	30.08	19.23	9.361***	2.905**
	(163.4)	(136.9)	(1.137)	(1.219)
-6	0.0156**	0.00396	0.00267	-0.00375**
	(0.00680)	(0.00628)	(0.00332)	(0.00188)
-5	-0.00745	0.0135	-0.000296	-0.00405
	(0.00980)	(0.0114)	(0.00330)	(0.00250)
-4	-0.00947	0.0182	0.000459	-0.00366
	(0.0150)	(0.0175)	(0.00369)	(0.00339)
-3	-0.0372*	0.0238	0.00139	-0.00280
	(0.0203)	(0.0234)	(0.00388)	(0.00369)
-2	-0.0640**	0.0317	0.00257	-0.00183
	(0.0262)	(0.0298)	(0.00412)	(0.00408)
-1	-0.0934***	0.0369	0.00389	-0.00102
	(0.0324)	(0.0363)	(0.00442)	(0.00452)
Constant	111.7	-77.19	-275.7***	-88.15**
	(895.3)	(.)	(33.00)	(35.38)
N	1468066	1468066	1468066	1468066
R^2	0.465	0.475	0.465	0.475
Mean coverage per 1000	6.98	6.98	34.21	34.21
Standard deviation	8.18	8.18	28.58	28.58

Notes: All models include district fixed effects, year fixed effects, district-year trends, and birth-year fixed effects. Covariates - parental education, age, age^2, and gender. Standard errors clustered by district and birth year in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 7: Dynamic treatment effects (Wooldridge, 2021)

	Binary Treatment		Continuous Treatment	
	(1) Math	(2) Literacy	(3) Math	(4) Literacy
TWFE × 2014 × 2014	0.0317* (0.0165)	0.0520** (0.0229)		
TWFE × 2014 × 2016	0.0937*** (0.0248)	0.0914*** (0.0266)		
TWFE × 2014 × 2018	0.182*** (0.0403)	0.138*** (0.0437)		
TWFE × 2016 × 2016	0.0247 (0.0343)	0.0395 (0.0382)		
TWFE × 2016 × 2018	0.0413 (0.0592)	0.0115 (0.0645)		
Coverage per 1000× 2014 × 2014			0.000138 (0.000481)	-0.000557 (0.000473)
Coverage per 1000× 2014 × 2016			0.00149*** (0.000483)	0.000745 (0.000481)
Coverage per 1000× 2014 × 2018			0.00203*** (0.000291)	0.000589** (0.000295)
Coverage per 1000× 2016 × 2016			0.00196 (0.00301)	0.00371 (0.00357)
Coverage per 1000× 2016 × 2018			0.000457 (0.00115)	-0.000338 (0.00127)
Constant	37.06 (38.02)	0.454 (39.28)	25.95 (38.21)	-2.766 (39.41)
N	2456787	2456787	2456787	2456787
R^2	0.401	0.423	0.401	0.423

Notes: All models include district fixed effects, year fixed effects, district-year trends, and birth-year fixed effects. Covariates - parental education, age, age^2, and gender. Standard errors clustered by district and birth year in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 8: Comparison of treatment effects

	Binary Treatment		Continuous Treatment	
Panel A: TWFE	(1) Math	(2) Literacy	(3) Math	(4) Literacy
TWFE	0.0377*** (0.0119)	0.0273** (0.0118)		
Coverage per 1000			0.00128*** (0.000171)	0.000115 (0.000169)
Panel B: Wooldridge (2021)				
TWFE	0.0379*** (0.00700)	0.0348*** (0.00762)		
Coverage per 1000			0.000470*** (0.000113)	0.000131 (0.000116)

Notes: All models include district fixed effects, year fixed effects, district-year trends, and birth-year fixed effects. Covariates - parental education, age, age^2, and gender. Standard errors clustered by district and birth year in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 9: Potential mechanisms of impact

	Binary Treatment			Continuous Treatment		
	(1) Toilet	(2) School Att.	(3) Thatched	(4) Toilet	(5) School Att.	(6) Thatched
TWFE	0.0187*** (0.00702)	0.00128 (0.00167)	0.00119 (0.00363)			
Coverage per 1000				0.00149*** (0.0000913)	0.0000593* (0.0000345)	-0.000312*** (0.0000713)
Constant	-33.04*** (11.64)	8.862** (4.213)	-8.766 (6.635)	-40.81*** (11.45)	8.551** (4.217)	-7.144 (6.626)
N	2456787	2456787	2456787	2456787	2456787	2456787
R^2	0.288	0.0725	0.251	0.289	0.0726	0.251

Notes: All models include district fixed effects, year fixed effects, district-year trends, and birth-year fixed effects. Covariates - parental education, age, age^2, and gender. Standard errors clustered by district and birth year in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 10: Effect on village-level outcomes

	Binary Treatment			Continuous Treatment		
	(1) Health Center	(2) Bank	(3) Post office	(4) Health Center	(5) Bank	(6) Post office
TWFE	0.0272 (0.0169)	-0.0154 (0.0130)	0.00977 (0.0204)			
Coverage per 1000				0.0000736 (0.000359)	-0.000267 (0.000296)	-0.0000378 (0.000362)
Constant	-36.89*** (1.256)	81.20*** (1.062)	74.53*** (1.221)	-37.33*** (2.280)	82.63*** (1.738)	74.70*** (2.145)
N	2456787	2456787	2456787	2456787	2456787	2456787
R^2	0.124	0.154	0.165	0.124	0.154	0.165

Notes: All models include district fixed effects, year fixed effects, and district-year trends. Covariates - parental education, age, age^2, and gender. Standard errors clustered by district in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 11: Determinants of program implementation

	(1)	(2)	(3)	(4)
Panel A: Data from GoI				
BPL households without toilet	-0.0150		-0.00763	
	(0.0162)		(0.0180)	
APL households without toilet		0.0150		0.00763
		(0.0162)		(0.0180)
Households without toilet	0.0393	0.0243	0.0439	0.0363
	(0.0240)	(0.0176)	(0.0264)	(0.0240)
Panel B: Data from ASER				
Percent of villages with post office	0.0192	0.0192	0.0141	0.0141
	(0.0159)	(0.0159)	(0.0204)	(0.0204)
Percent of villages with bank	-0.0326	-0.0326	-0.0401	-0.0401
	(0.0309)	(0.0309)	(0.0306)	(0.0306)
Percent of villages with primary health center	0.0314	0.0314	0.0296	0.0296
	(0.0272)	(0.0272)	(0.0280)	(0.0280)
Percent of thatched houses			-0.00894	-0.00894
			(0.0256)	(0.0256)
Percent of houses with toilet			0.0220	0.0220
			(0.0212)	(0.0212)
Percent of school enrolled children			-0.0306	-0.0306
			(0.132)	(0.132)
Mean age of children			1.491^{*}	1.491^{*}
			(0.866)	(0.866)
Percent of male children			-0.106	-0.106
			(0.0841)	(0.0841)
Percent of mothers attended school			0.0254	0.0254
			(0.0517)	(0.0517)
Percent of fathers attended school			-0.0115	-0.0115
			(0.0615)	(0.0615)
Constant	1157.7^{***}	1157.7^{***}	1393.0^{**}	1393.0^{**}
	(195.9)	(195.9)	(523.8)	(523.8)
N	2177	2177	2177	2177
R^2	0.0901	0.0901	0.0955	0.0955

Notes: GoI is an acronym for Government of India. APL and BPL are acronyms for above and below poverty lines. Data are aggregated at the district level. All regressions include state fixed effects, year fixed effects, state-year trends. Standard errors clustered by state in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 12: Raw and binary math and literacy outcomes

	Binary Treatment		Continuous Treatment	
Panel A: Raw scores	(1) Math	(2) Literacy	(3) Math	(4) Literacy
TWFE	0.0463***	0.0363**		
	(0.0148)	(0.0152)		
Coverage per 1000			0.00159***	0.0000693
			(0.000223)	(0.000245)
Constant	48.75	8.006	40.36	7.574
	(29.70)	(28.62)	(29.77)	(28.67)
N	2456787	2456787	2456787	2456787
R^2	0.478	0.491	0.478	0.491

Panel B: Binary scores	(1) Math	(2) Literacy	(3) Math	(4) Literacy
TWFE	0.0179***	0.0179***		
	(0.00647)	(0.00633)		
Coverage per 1000			0.000570***	0.0000281
			(0.0000951)	(0.0000983)
Constant	15.96	-9.984	12.96	-10.16
	(14.58)	(13.72)	(14.61)	(13.73)
N	2456787	2456787	2456787	2456787
R^2	0.270	0.361	0.270	0.361

Notes: In Panel B, math equals 1 if the child can do three by one digit division (has level 4 proficiency) and it equals 0 for the other proficiency levels. Literacy equals 1 if the child has a level 4 proficiency, i.e., can read grade 2 text, and 0 for the other levels. All regressions include district fixed effects, year fixed effects, birth-year fixed effects, and district-year trends. Standard errors clustered by district and birth year in parentheses. *** p<0.01, ** p<0.05, * p<0.1

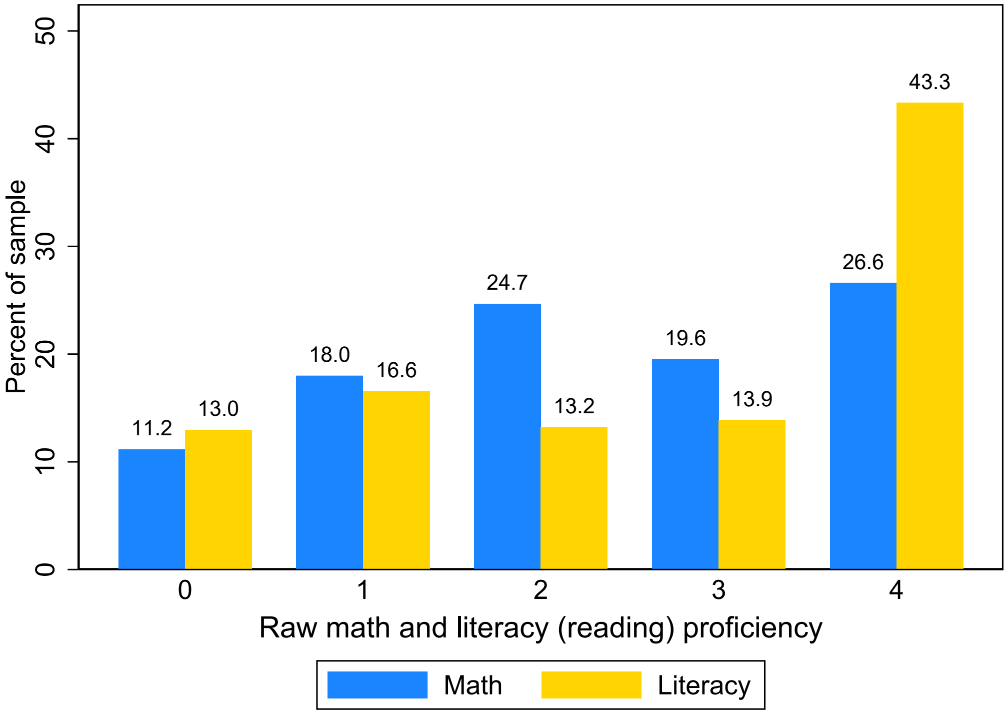
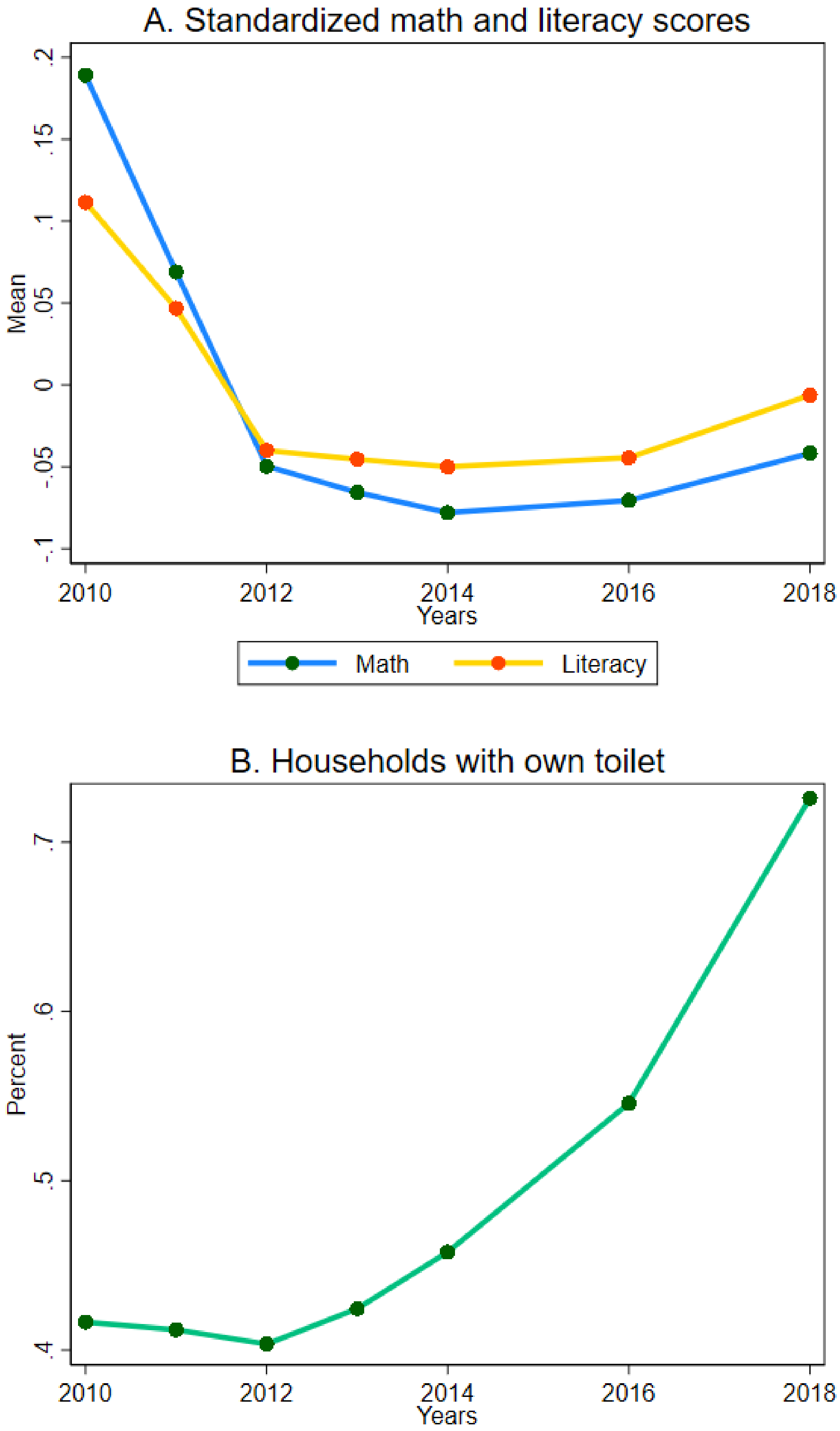


Figure 1: Math and literacy proficiency rates of sample children


A. Standardized math and literacy scores

B. Households with own toilet

Figure 2: Standardized math and literacy scores and percentage of sample with household toilet in ASER

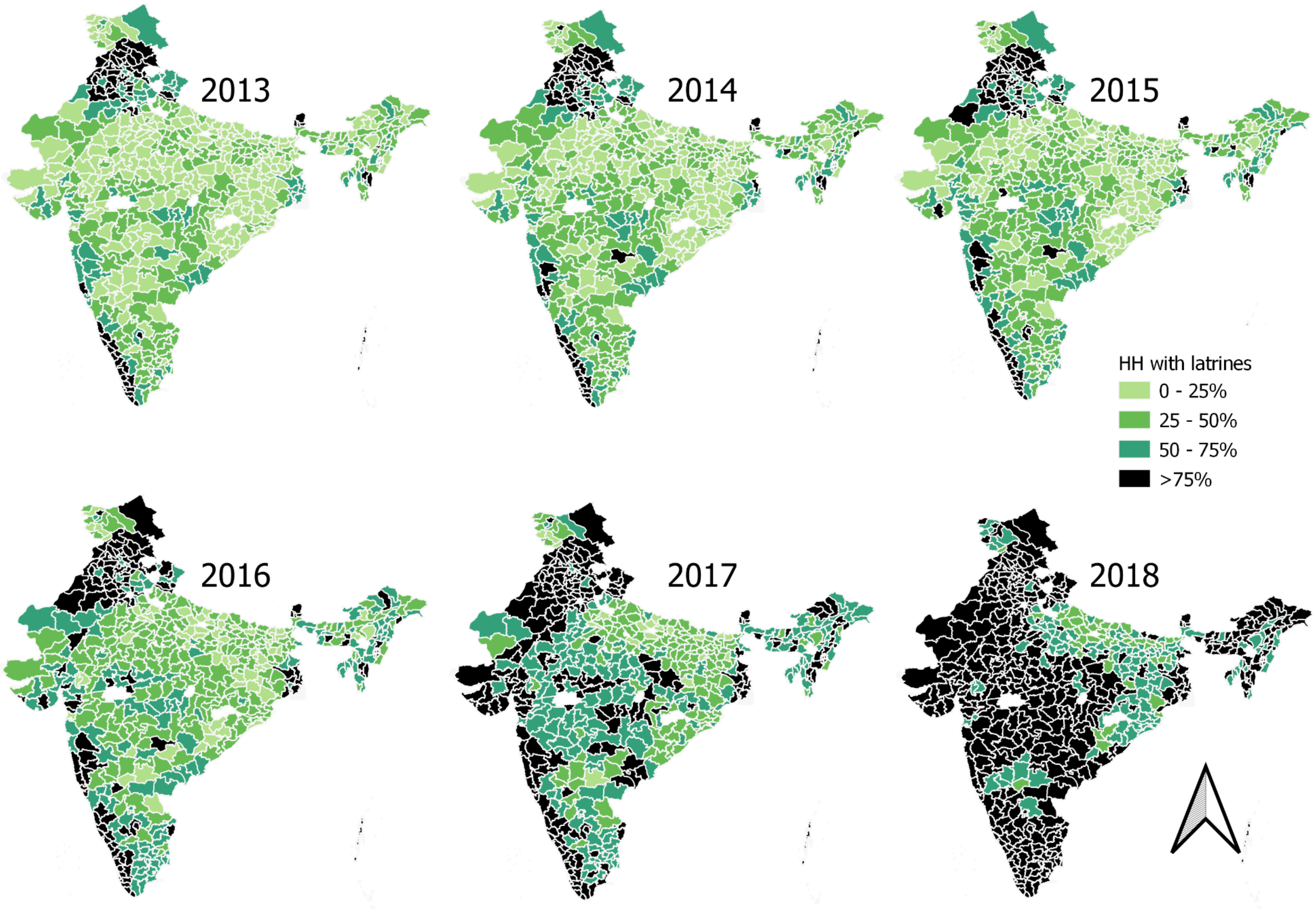


Figure 3: Percentage of households with own toilets (latrines) in a district (calculated using the CIM coverage data)

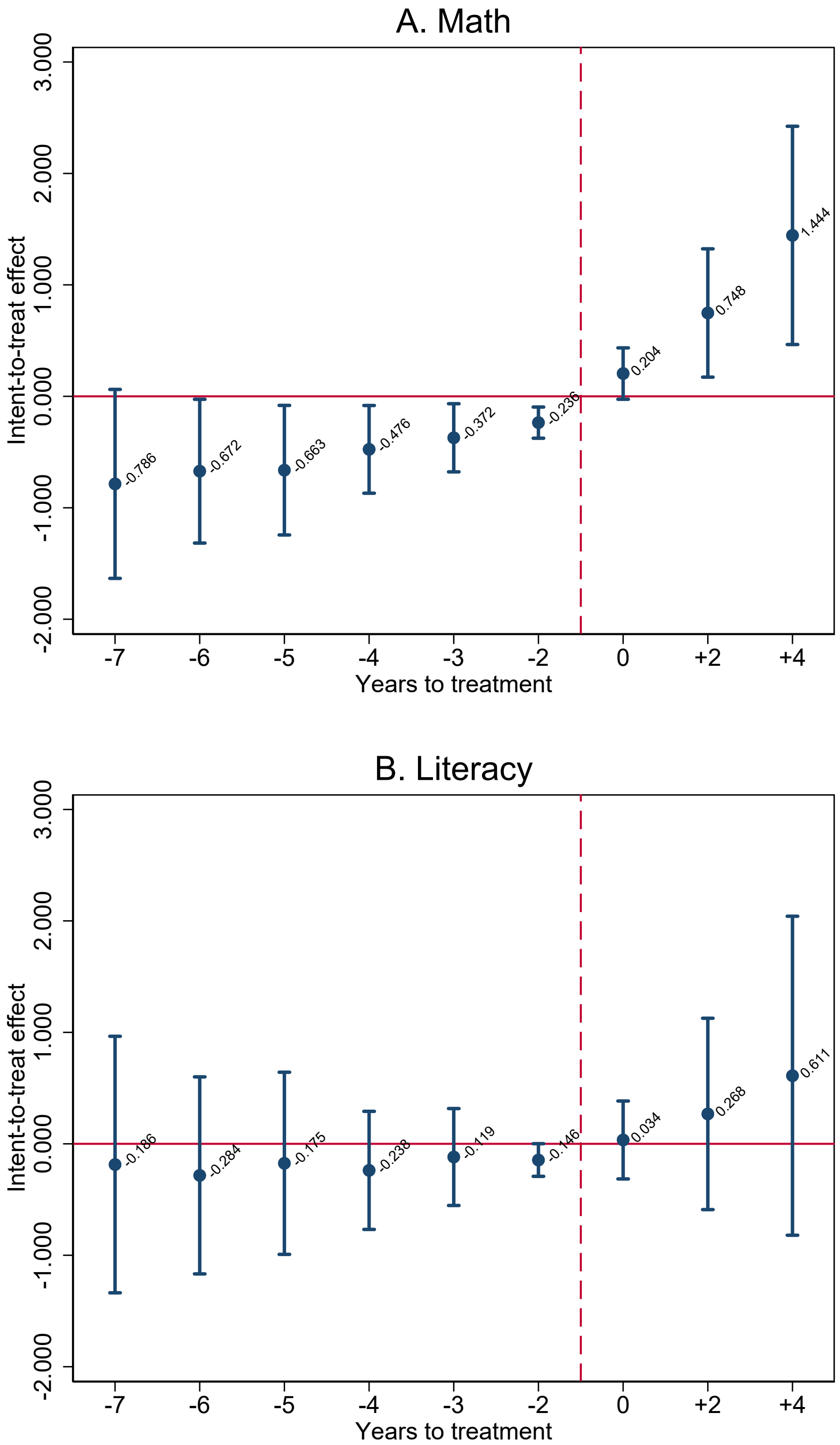


Figure 4: Event study

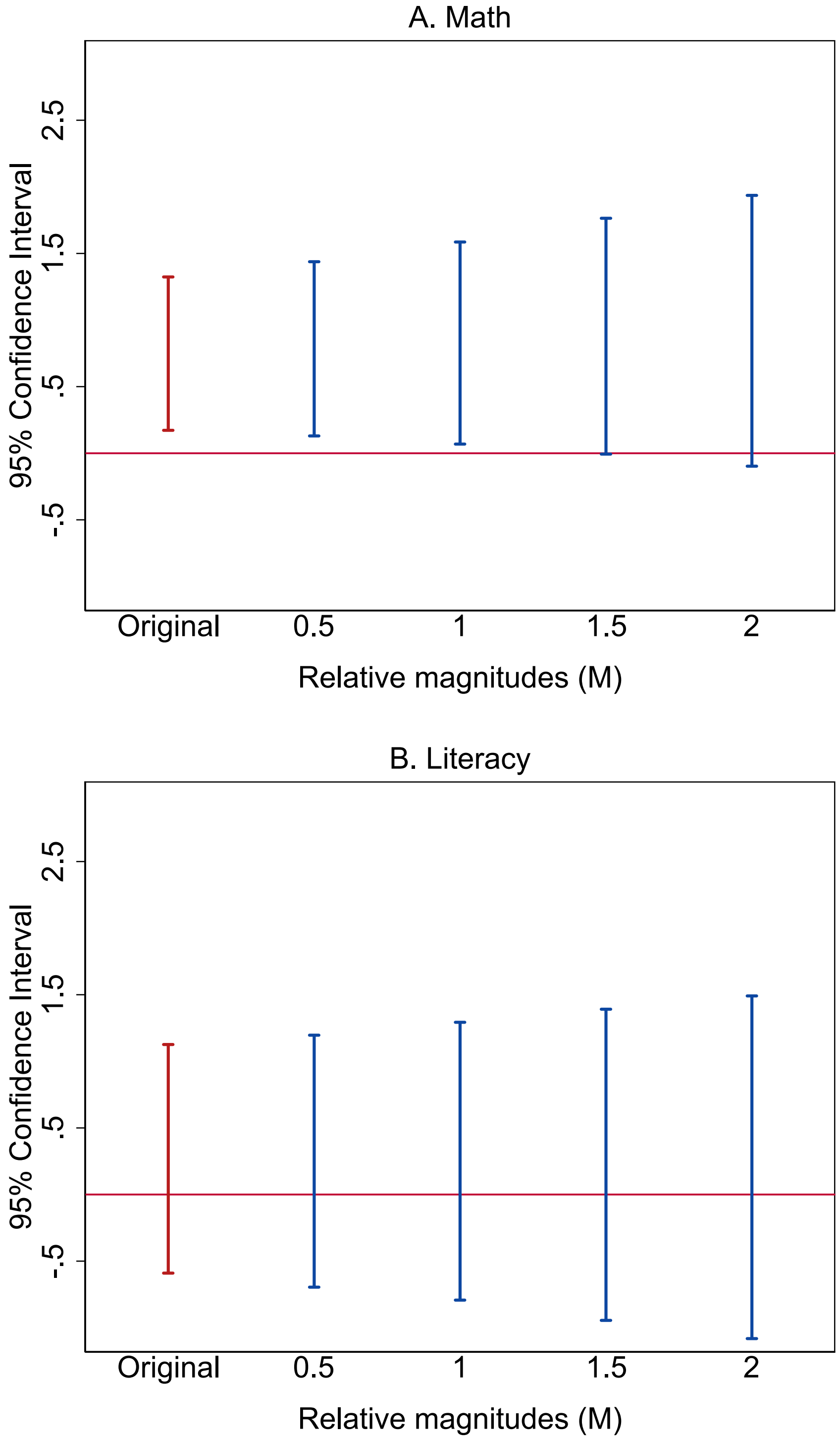


Figure 5: Sensitivity of 2016 event study estimate using relative magnitudes (Rambachan and Roth, 2023)

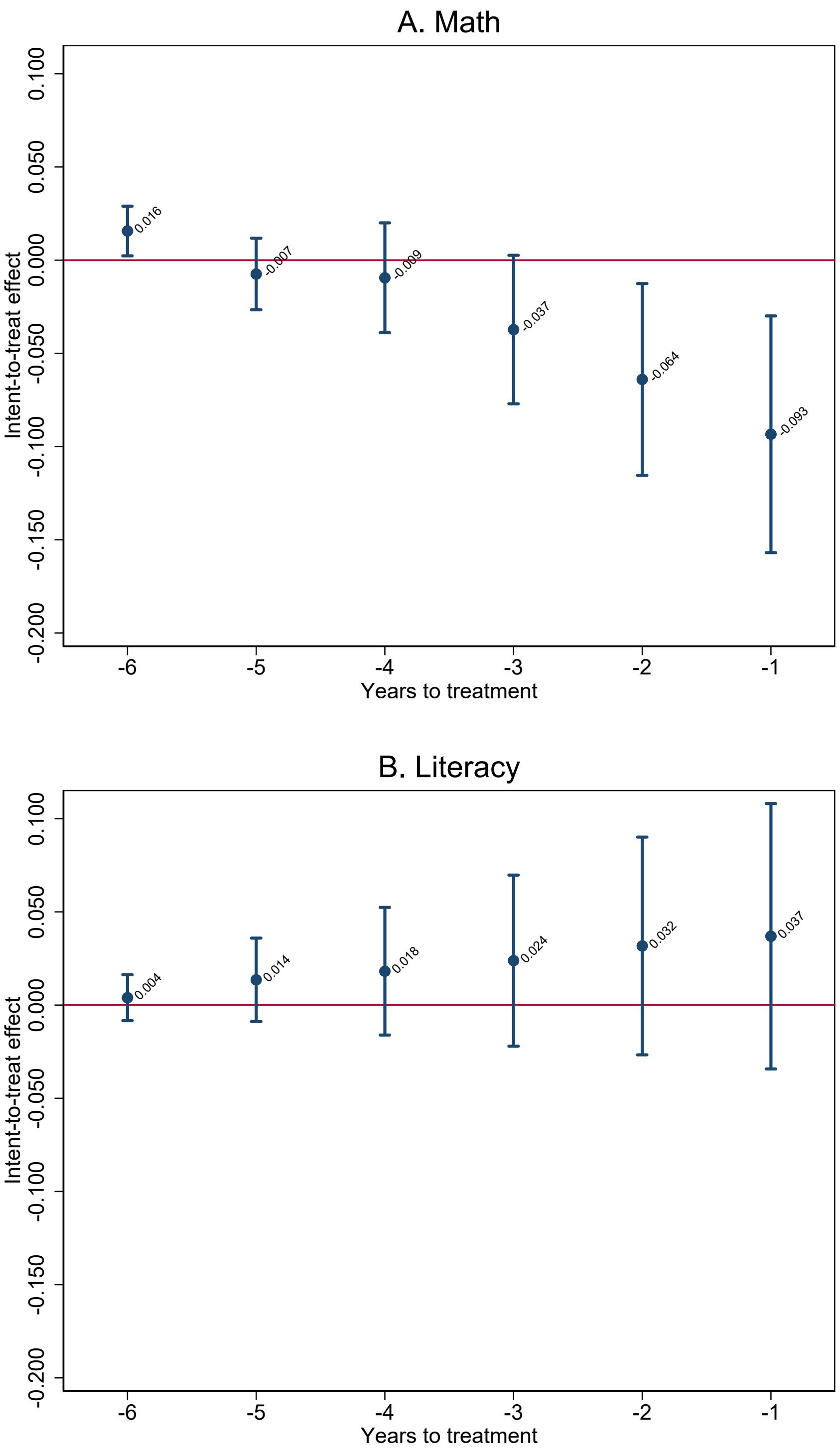


Figure 6: Effect of first year’s coverage on pre-treatment outcomes

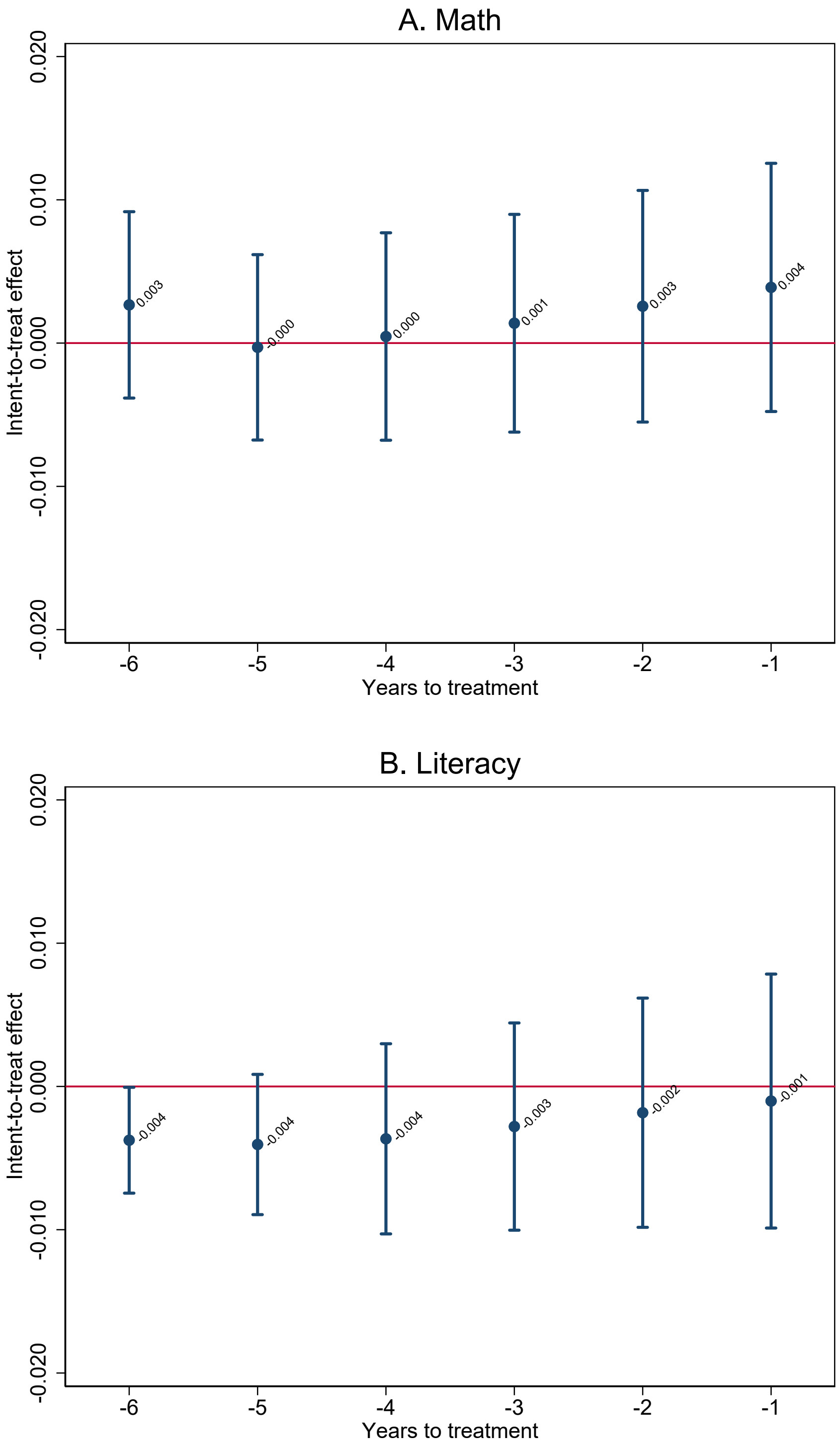


Figure 7: Effect of most recent year's coverage on pre-treatment outcomes

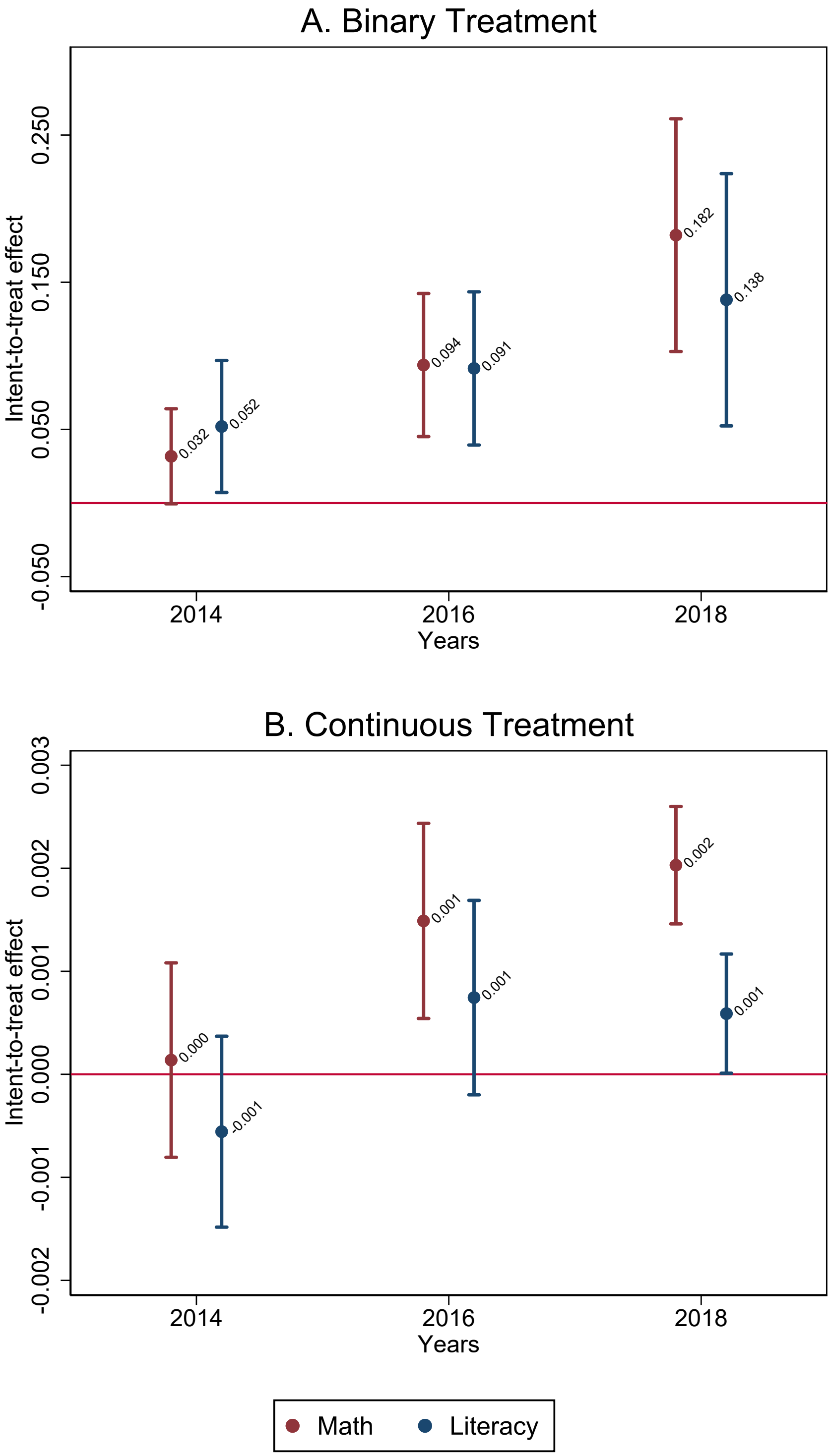


Figure 8: Dynamic treatment effect estimates of Math and Literacy (Wooldridge, 2021)

Appendix

Table A.1: Average characteristics across ASER and CIM rounds (balanced sample)

	2010	2011	2012	2013	2014	2016	2018
	N=363,275	N=344,406	N=310,525	N=310,981	N=298,390	N=281,471	N=276,050
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Math	0.192	0.069	-0.055	-0.069	-0.078	-0.073	-0.041
	(1.003)	(1.000)	(0.992)	(0.998)	(0.998)	(0.990)	(0.981)
Literacy	0.115	0.044	-0.044	-0.049	-0.048	-0.045	-0.004
	(0.933)	(0.963)	(1.006)	(1.019)	(1.035)	(1.030)	(1.016)
Age in years	10.384	10.407	10.315	10.281	10.293	10.298	10.232
	(3.234)	(3.253)	(3.257)	(3.266)	(3.258)	(3.281)	(3.264)
Male (1=Yes, 0=No)	0.547	0.530	0.517	0.513	0.509	0.501	0.501
	(0.498)	(0.499)	(0.500)	(0.500)	(0.500)	(0.500)	(0.500)
Father attended school (1=Yes, 0=No)	0.719	0.729	0.733	0.733	0.743	0.744	0.765
	(0.450)	(0.444)	(0.443)	(0.442)	(0.437)	(0.436)	(0.424)
Mother attended school (1=Yes, 0=No)	0.509	0.503	0.500	0.518	0.529	0.557	0.594
	(0.500)	(0.500)	(0.500)	(0.500)	(0.499)	(0.497)	(0.491)
Thatched house (1=Yes, 0=No)	0.351	0.350	0.350	0.334	0.315	0.300	0.260
	(0.477)	(0.477)	(0.477)	(0.472)	(0.464)	(0.458)	(0.438)
Toilet in house (1=Yes, 0=No)	0.400	0.395	0.386	0.413	0.440	0.534	0.716
	(0.490)	(0.489)	(0.487)	(0.492)	(0.496)	(0.499)	(0.451)
Attends school (1=Yes, 0=No)	0.962	0.962	0.967	0.971	0.971	0.971	0.974
	(0.191)	(0.191)	(0.179)	(0.168)	(0.168)	(0.168)	(0.159)
Health center in village (1=Yes, 0=No)	0.416	0.425	0.429	0.422	0.419	0.421	0.405
	(0.493)	(0.494)	(0.495)	(0.494)	(0.493)	(0.494)	(0.491)
Bank in village (1=Yes, 0=No)	0.219	0.234	0.244	0.253	0.262	0.296	0.300
	(0.414)	(0.423)	(0.429)	(0.435)	(0.440)	(0.457)	(0.458)
Post office in village (1=Yes, 0=No)	0.447	0.444	0.444	0.420	0.419	0.410	0.388
	(0.497)	(0.497)	(0.497)	(0.494)	(0.493)	(0.492)	(0.487)
CIM coverage in district in current year					8,989.545	20,386.057	52,254.154
					(10,450.804)	(23,144.465)	(48,735.592)
CIM coverage in district in previous year						9,577.407	34,582.073
						(14,374.786)	(37,884.524)
CIM coverage per 1000 in current year					6.549	15.614	33.893
					(8.331)	(14.624)	(28.359)
CIM coverage per 1000 in previous year						6.805	24.981
						(7.777)	(21.615)
Percent rural households with own toilet				34.961	38.138	48.960	79.398
				(25.209)	(25.894)	(27.207)	(36.383)
Percent rural population with own toilet				6.962	7.554	9.810	15.626
				(5.159)	(5.297)	(5.546)	(6.346)

Notes: CIM coverage means the number of toilets constructed under the program in the district. CIM coverage in previous year for 2018, and 2016 are for 2017, and 2015. Data obtained from: Website

Table A.2: Difference in average characteristics between full and balanced sample of districts

	Full sample N=2,456,787 (1)	Balanced sample N=2,185,098 (2)	Omitted sample N=271,689 (3)	Diff. (1) - (2) (4)	Diff. (1) - (3) (5)
Math	-0.000	-0.000	-0.000	-0.000	0.000
	(1.000)	(1.000)	(1.000)	(0.001)	(0.002)
Literacy	-0.000	-0.000	-0.000	-0.000	0.000
	(1.000)	(1.000)	(1.000)	(0.001)	(0.002)
Age in years	10.319	10.321	10.310	-0.001	0.010
	(3.262)	(3.259)	(3.290)	(0.003)	(0.007)
Male (1=Yes, 0=No)	0.517	0.518	0.510	-0.001*	0.007***
	(0.500)	(0.500)	(0.500)	(0.000)	(0.001)
Father attended school (1=Yes, 0=No)	0.738	0.737	0.745	0.001**	-0.007***
	(0.440)	(0.440)	(0.436)	(0.000)	(0.001)
Mother attended school (1=Yes, 0=No)	0.532	0.528	0.565	0.004***	-0.033***
	(0.499)	(0.499)	(0.496)	(0.000)	(0.001)
Thatched house (1=Yes, 0=No)	0.326	0.325	0.335	0.001**	-0.008***
	(0.469)	(0.468)	(0.472)	(0.000)	(0.001)
Toilet in house (1=Yes, 0=No)	0.479	0.462	0.615	0.017***	-0.137***
	(0.500)	(0.499)	(0.486)	(0.000)	(0.001)
Attends school (1=Yes, 0=No)	0.969	0.968	0.978	0.001***	-0.009***
	(0.173)	(0.176)	(0.146)	(0.000)	(0.000)
Health center in village (1=Yes, 0=No)	0.426	0.420	0.479	0.007***	-0.052***
	(0.495)	(0.494)	(0.500)	(0.000)	(0.001)
Bank in village (1=Yes, 0=No)	0.257	0.256	0.268	0.001***	-0.011***
	(0.437)	(0.436)	(0.443)	(0.000)	(0.001)
Post office in village (1=Yes, 0=No)	0.424	0.426	0.410	-0.002***	0.015***
	(0.494)	(0.495)	(0.492)	(0.000)	(0.001)
CIM coverage in district in current year	25,303.324	26,691.132	16,359.418	-1,387.808***	8,943.906***
	(35,223.242)	(36,240.061)	(26,076.343)	(52.705)	(100.147)
CIM coverage in district in previous year	20,833.960	21,958.175	14,209.566	-1,124.215***	6,624.394***
	(30,547.128)	(31,164.965)	(25,634.674)	(56.241)	(104.259)
CIM coverage per 1000 in current year	18.834	18.349	21.955	0.484***	-3.121***
	(23.012)	(21.974)	(28.625)	(0.033)	(0.069)
CIM coverage per 1000 in previous year	16.121	15.805	17.982	0.316***	-1.862***
	(18.712)	(18.559)	(19.484)	(0.034)	(0.065)
Percent rural households with own toilet	50.729	49.662	58.388	1.067***	-7.646***
	(34.074)	(33.665)	(35.961)	(0.043)	(0.090)
Percent rural population with own toilet	10.087	9.850	11.790	0.237***	-1.700***
	(6.819)	(6.532)	(8.406)	(0.008)	(0.018)

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

Table A.3: Two way fixed effects (balanced sample)

					Early v. Late/No	
	(1) Math-Z	(2) Literacy	(3) Math	(4) Literacy	(5) Math	(6) Literacy
TWFE	0.0435**	0.0343	0.0477***	0.0379***	0.0482***	0.0372***
	(0.0211)	(0.0220)	(0.0128)	(0.0126)	(0.0132)	(0.0128)
Constant	79.44**	-21.28	84.63***	-16.17	84.63***	-16.17
	(39.86)	(42.34)	(26.41)	(22.50)	(26.41)	(22.50)
N	2185098	2185098	2185098	2185098	2185098	2185098
R^2	0.394	0.415	0.449	0.467	0.449	0.467
Covariates			Yes	Yes	Yes	Yes

Notes: All models include district fixed effects, year fixed effects, district-year trends, and birth-year fixed effects. Covariates - parental education, age, age^2, and gender. Standard errors clustered by district and birth year in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A.4: Continuous coverage (balanced sample)

	(1) Math	(2) Literacy
Coverage per 1000	0.00122***	0.0000791
	(0.000176)	(0.000175)
Constant	81.04***	-16.40
	(26.44)	(22.52)
N	2185098	2185098
R^2	0.449	0.467
Mean coverage per 1000	12.58	12.58
Standard deviation	16.37	16.37

Notes: Both models include district fixed effects, year fixed effects, district-year trends, and birth-year fixed effects. Covariates - parental education, age, age^2, and gender. Standard errors clustered by district and birth year in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A.5: Quadratic of continuous coverage

Panel A: Full sample	(1) Math	(2) Literacy
Coverage per 1000	0.00161***	-0.000388
	(0.000343)	(0.000343)
Coverage per 1000^2	-0.00000441	0.00000673**
	(0.00000338)	(0.00000343)
Constant	32.98	5.111
	(25.87)	(20.45)
N	2456787	2456787
R^2	0.449	0.467

Panel B: Balanced sample	(1) Math	(2) Literacy
Coverage per 1000	0.00164***	-0.000376
	(0.000356)	(0.000357)
Coverage per 1000^2	-0.00000557	0.00000602*
	(0.00000345)	(0.00000351)
Constant	80.66***	-16.00
	(26.45)	(22.52)
N	2185098	2185098
R^2	0.449	0.467

Notes: Both models include district fixed effects, year fixed effects, district-year trends, and birth-year fixed effects. Covariates - parental education, age, age^2, and gender. Standard errors clustered by district and birth year in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table A.6: Heterogeneity across gender and age (balanced sample)

Panel A: Gender	Math		Literacy		Math		Literacy	
	(1) Boys	(2) Girls	(3) Boys	(4) Girls	(5) Boys	(6) Girls	(7) Boys	(8) Girls
TWFE	0.0477***	0.0467***	0.0411***	0.0312**				
	(0.0137)	(0.0166)	(0.0152)	(0.0152)				
Coverage per 1000					0.00143***	0.001000***	0.000193	-0.0000228
					(0.000216)	(0.000216)	(0.000219)	(0.000207)
Constant	81.28***	87.77**	-10.64	-22.79	77.01***	84.85**	-11.22	-22.72
	(25.74)	(34.81)	(26.12)	(27.82)	(25.77)	(34.83)	(26.13)	(27.83)
N	1132092	1053006	1132092	1053006	1132092	1053006	1132092	1053006
R^2	0.465	0.436	0.475	0.463	0.465	0.436	0.475	0.463
Mean coverage per 1000					12.49	12.67	12.49	12.67
Standard deviation					16.27	16.48	16.27	16.48

Panel B: Age groups	(1) 5 - 10 years	(2) 11 - 16 years	(3) 5 - 10 years	(4) 11 - 16 years	(5) 5 - 10 years	(6) 11 - 16 years	(7) 5 - 10 years	(8) 11 - 16 years
TWFE	0.0446***	0.0530***	0.0409**	0.0345***				
	(0.0168)	(0.0164)	(0.0188)	(0.0129)				
Coverage per 1000					0.000676***	0.00176***	-0.0000861	0.000298
					(0.000232)	(0.000234)	(0.000259)	(0.000200)
Constant	57.64	116.6***	-30.84	3.435	55.59	111.5***	-30.58	2.582
	(39.54)	(17.47)	(37.28)	(14.02)	(39.61)	(17.36)	(37.30)	(14.01)
N	1156617	1028481	1156617	1028481	1156617	1028481	1156617	1028481
R^2	0.348	0.112	0.349	0.0823	0.348	0.112	0.349	0.0823
Mean coverage per 1000					12.52	12.64	12.52	12.64
Standard deviation					16.44	16.29	16.44	16.29

Notes: All models include district fixed effects, year fixed effects, district-year trends, and birth-year fixed effects. Covariates - parental education, age, age^2, and gender. Standard errors clustered by district and birth year in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A.7: Event study (balanced sample)

	(1) Math	(2) Literacy
-7	-0.412	0.345
	(0.439)	(0.577)
-6	-0.466	0.142
	(0.347)	(0.438)
-5	-0.542*	0.0842
	(0.304)	(0.408)
-4	-0.318	0.0662
	(0.212)	(0.261)
-3	-0.303*	0.0200
	(0.160)	(0.217)
-2	-0.112	0.0151
	(0.0736)	(0.0670)
0	0.277**	0.0636
	(0.115)	(0.179)
+2	0.751**	0.139
	(0.296)	(0.434)
+4	1.362***	0.295
	(0.504)	(0.719)
Constant	84.96***	-16.23
	(26.42)	(22.50)
N	2185098	2185098
R^2	0.449	0.467

Notes: Both models include district fixed effects, year fixed effects, district-year trends, and birth-year fixed effects. Covariates - parental education, age, age^2, and gender. Standard errors clustered by district and birth year in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A.8: Upper and lower bounds for treatment effect estimate of period +2 (Rambachan and Roth, 2023)

	Math		Literacy	
Panel A: Relative magnitudes (M)	(1) Lower bound	(2) Upper bound	(3) Lower bound	(4) Upper bound
Original	0.172	1.324	-0.590	1.126
M=0.5	0.130	1.438	-0.695	1.196
M=1.0	0.069	1.586	-0.793	1.293
M=1.5	-0.007	1.764	-0.945	1.391
M=2.0	-0.097	1.936	-1.082	1.491

Panel B: Smoothness bounds (N)	(1) Lower bound	(2) Upper bound	(3) Lower bound	(4) Upper bound
Original	0.172	1.324	-0.590	1.126
N=0.01	-0.107	0.472	0.083	0.692
N=0.02	0.023	0.700	0.127	0.828
N=0.03	0.018	0.731	0.097	0.851
N=0.04	-0.000	0.755	0.057	0.872
N=0.05	-0.019	0.779	0.016	0.894

Table A.9: Dynamic treatment effects with controls (Wooldridge, 2021)

	Binary Treatment		Continuous Treatment	
	(1) Math	(2) Literacy	(3) Math	(4) Literacy
TWFE × DIST14 × F14	0.0353***	0.0560***		
	(0.0121)	(0.0187)		
TWFE × DIST14 × F16	0.0932***	0.0927***		
	(0.0191)	(0.0202)		
TWFE × DIST14 × F18	0.194***	0.149***		
	(0.0265)	(0.0294)		
TWFE × DIST16 × F16	0.0263	0.0408		
	(0.0233)	(0.0275)		
TWFE × DIST16 × F18	0.0664*	0.0320		
	(0.0353)	(0.0392)		
Coverage per 1000 × DIST14 × F14			0.000265	-0.000437
			(0.000344)	(0.000333)
Coverage per 1000 × DIST14 × F16			0.00142***	0.000701**
			(0.000339)	(0.000342)
Coverage per 1000 × DIST14 × F18			0.00191***	0.000480***
			(0.000175)	(0.000167)
Coverage per 1000 × DIST16 × F16			0.00167	0.00339
			(0.00190)	(0.00221)
Coverage per 1000 × DIST16 × F18			0.000530	-0.000262
			(0.000612)	(0.000693)
Constant	44.82*	9.585	34.33	6.953
	(25.33)	(20.77)	(25.42)	(20.86)
N	2456787	2456787	2456787	2456787
R^2	0.452	0.471	0.452	0.471

Notes: All models include district fixed effects, year fixed effects, district-year trends, and birth-year fixed effects. Covariates - parental education, age, age^2, and gender. Standard errors clustered by district and birth year in parentheses. *** p<0.01, ** p<0.05, * p<0.1

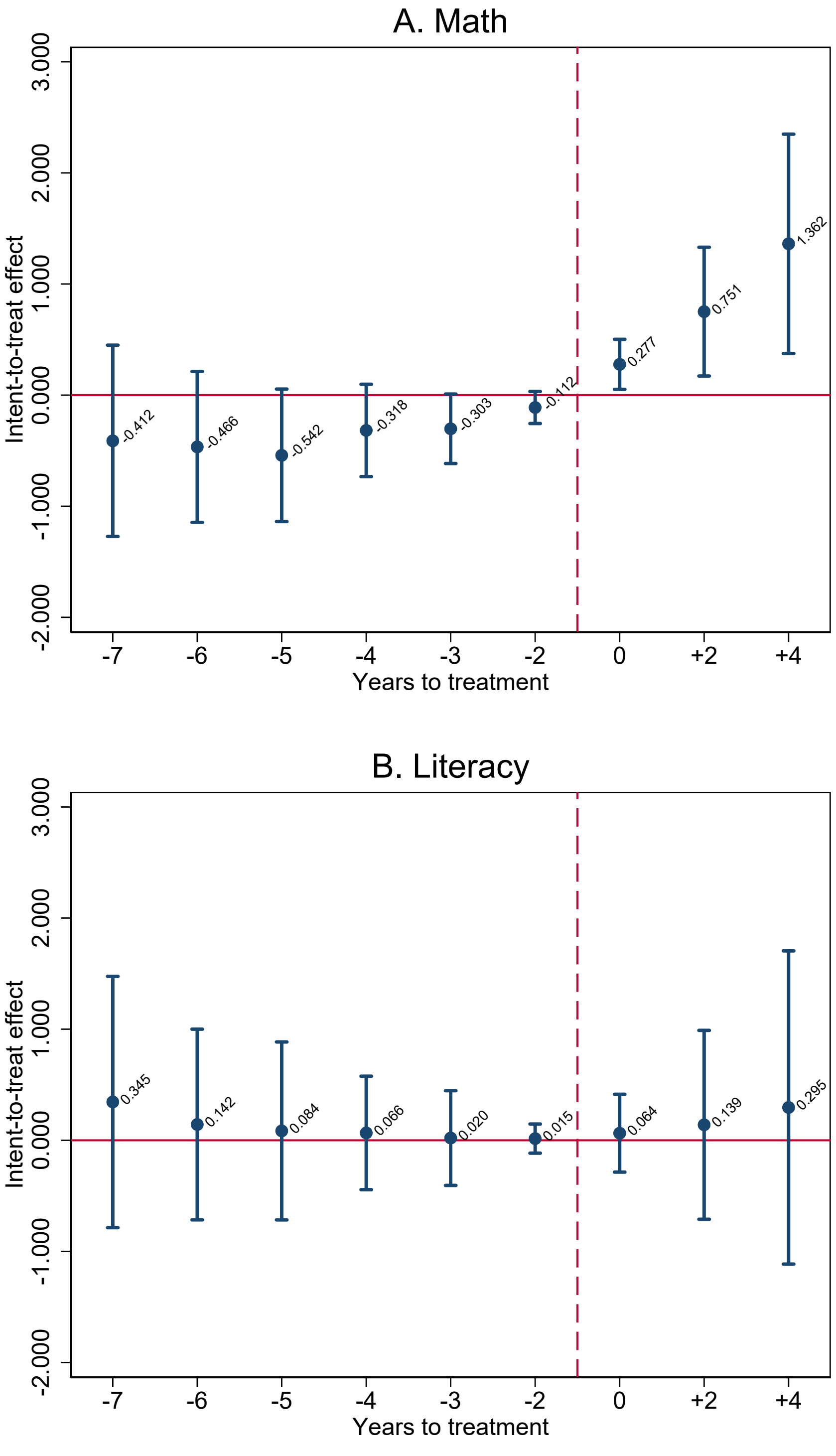


Figure A.1: Event study (balanced sample)

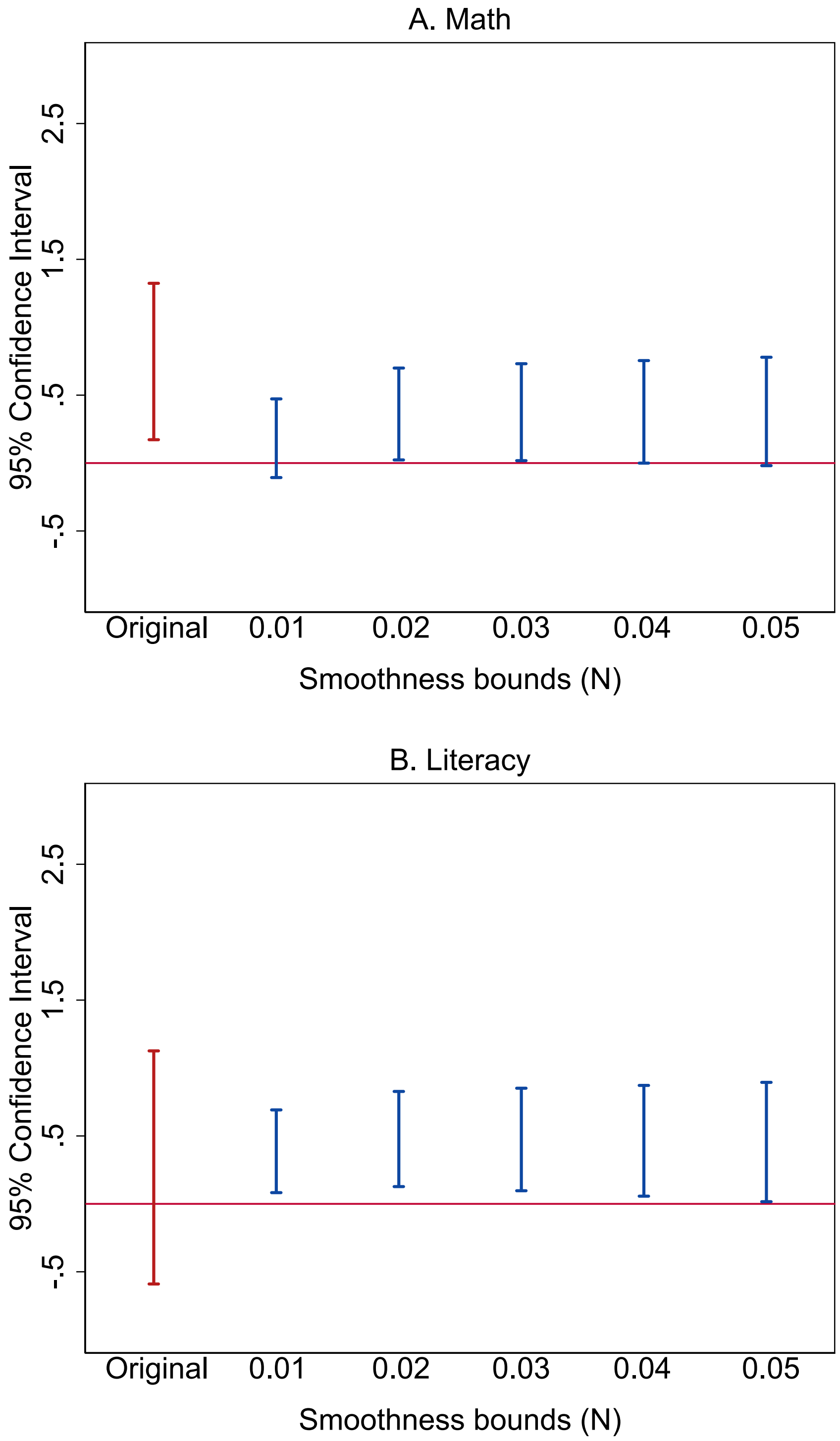


Figure A.2: Sensitivity of 2016 event study estimate using smoothness bounds (Rambachan and Roth, 2023)

Violation of pre-treatment trends. Smoothness bounds method

In addition to the Clean India Mission, the newly elected government implemented multiple programs which can potentially confound the treatment effects we estimate in the TWFE, and event study regressions. These concurrent programs (e.g., Sansad Adarsh Gram Yojana, Saksham scheme, People's Wealth Scheme, and Jeevan Jyoti Bima Yojana) can affect children's cognition and lead to large differences between those in treatment and control districts. Therefore, as a result of these programs, it is reasonable to assume that pre-treatment differential trends (pre-treatment differences between children in treatment and control districts) can evolve randomly and vary between the pre-treatment and post-treatment periods. Based on this fact, we adopt the second method proposed by Rambachan and Roth (2023) for testing the sensitivity of the intent-to-treat effect estimate for period +2.

In this approach, we introduce a smoothness restrictions parameter ($N \geq 0$). N represents the extent to which post-treatment differential trends (counterfactual) can differ from pre-treatment differential trends. Mathematically, this is represented as $Post_{trends} = N \pm Pre_{trends}$. N=0 implies that differential trends are linear, i.e., pre-existing cognition differences between children in treatment and control districts evolve smoothly over time. But due to other government programs, the differences in differential trends can deviate from zero. This means that N stands for the change in post-treatment differential trends from pre-treatment differential trends. Values of N we use are 0.01, 0.02, 0.03, 0.04, and 0.05.

We assume that changes in differential trends between post- and pre-treatment periods occur due to government programs (e.g., Sansad Adarsh Gram Yojana, Saksham scheme, People's Wealth Scheme, and Jeevan Jyoti Bima Yojana) implemented alongside CIM. The purpose of these programs is to improve household income, children's health, access to food, school enrollment, and parental labor force participation. All of these changes can affect children's cognition outcomes positively. To our knowledge, however, there currently exists limited research studying the effect of these programs. Therefore, to interpret the choices of N, we turn to the literature from years prior to CIM.

For example, Adukia et al. (2020) study the effect of a road construction program, the Pradhan Mantri Gram Sadak Yojana scheme (2020). They find that the probability of a village being connected to a paved road is associated with a 0.058 percentage point improvement in the log number of children who pass middle school exams in India. This means that a value of N=0.01, in our study, corresponds to 0.017 times the probability of a road being constructed in the village, through a similar program, around the implementation of CIM in the village. Similarly, Chakraborty and Jayaraman (2019) examine the effect of the Mid-day Meal program (1995) on children's cognition outcomes in India. Their findings show that increased exposure to free meals at government schools increases children's cognition scores by 0.03 percentage points. Therefore, N=0.01, in our analysis, corresponds to a 0.33 times improvement in children's exposure to meals/food from programs implemented around the Clean India Mission in 2014.

These results are shown in Table A.8 and plotted in Figure A.2. For math scores, shown in Panel A of the figure, confidence intervals (0.172 and 1.323) for the event study coefficient for period +2 is shown in red. Blue bars show the confidence intervals of the intent-to-treat estimate for period +2 using different values of N. We find that if the difference between post-treatment and pre-treatment differentials trend is restricted to 0.01, the intent-to-treat effect estimate for period +2 is positive but statistically insignificant. The lower and upper bounds for this coefficient are -0.109 and 0.472. Looking further to the right, we find that the upper and lower bounds for the post-treatment coefficients for period +2 are statistically significant for N=0.03 and 0.04. This means that we can reject the null effects of CIM on math scores (in period +2) until the effect of co-existing government programs is large enough to cause changes in pre-treatment differential trends in children's cognition by more than 0.04 SD. For literacy, the estimated confidence intervals, shown in Panel B of Figure A.2, are close to zero for all values of N. All the intent-to-treat effect estimate on literacy in our event study analysis are statistically insignificant. Hence, the estimated coefficients in Panel B of Figure A.2 should be interpreted with caution.