

Research Paper

Sanitation and externalities: evidence from early childhood health in rural India

Luis Andrés, Bertha Briceño, Claire Chase and Juan A. Echenique

ABSTRACT

This paper estimates two sources of benefits related to sanitation infrastructure access: a direct benefit households receive when they have access to sanitation infrastructure, and an external benefit produced by the neighborhood's access to sanitation infrastructure. Using a sample of children under age four from rural areas of India in the Third Round of District Level Household Survey 2007–08, the study demonstrates evidence of positive direct benefits and a concave positive externality for improved sanitation and fixed-point defecation. The paper finds that a child who moves from a household without improved sanitation and a low ratio of village access to a household with improved sanitation and a high ratio of village access enjoys a reduction in diarrhea prevalence of 47 percent. From this, one-fourth of this benefit is due to the direct benefit, leaving the rest to external gains. These results hold under several robustness checks.

Key words | diarrhea, externalities, health, India, sanitation

Luis Andrés (corresponding author)

Claire Chase

World Bank,
1818 H Street NW,
Washington, DC,
USA
E-mail: landres@worldbank.org

Bertha Briceño

Inter-American Development Bank,
1300 New York Avenue NW,
Washington, DC,
USA

Juan A. Echenique

University of Maryland,
Van Munching Hall,
College Park,
USA

INTRODUCTION

This paper investigates the existence and size of benefits from access to sanitation on child health outcomes in rural India. We emphasize in particular two sources of benefit. The first is a direct benefit, which is privately enjoyed by a household whose members stop defecating in the open, and move to either fixed-point defecation or an improved sanitation facility. (In this paper we refer to 'fixed-point defecation' as defecation into a pit or other containment structure, regardless of the quality of the structure or whether it is hygienically maintained (i.e., includes access to both improved and unimproved facilities). The WHO-UNICEF Joint Monitoring Program for Water Supply and Sanitation defines 'improved sanitation' as 'access to a sanitation facility that hygienically separates human excreta from human contact'. Open defecation is defined as, 'not having access to any type of toilet' and implies 'going in the field'. The second benefit is an external one, also called externality, generated by a neighbor's access to

sanitation that results in a lower probability of human contact with human excreta. In this paper we seek to answer two questions: (i) what is the private benefit of a household's access to sanitation infrastructure on early childhood health outcomes? and (ii) what is the benefit produced through externalities of a neighbor's access to this sanitation infrastructure? These questions are motivated by the poor sanitation situation worldwide, where, by the end of 2015, 2.4 billion people still lacked access to improved sanitation (JMP 2015). In India alone there are an estimated 62% of households without improved sanitation and 51% of households practice open defecation (International Institute for Population Sciences 2010).

There is growing literature on the benefits perceived by members of a household with access to improved sanitation. There is strong evidence from observational, quasi-experimental, and small-scale intervention studies that the individual benefit to having improved sanitation is

potentially very large. Evidence from experimental evidence (Patil *et al.* 2014) and reduced form estimations (Headey 2013; Duflo *et al.* 2015; Geruso & Spears 2015) indicates positive difference. But how much of this benefit is actually driven through a community benefit, or externality to private sanitation, is less well understood. Hence, there is a need to empirically quantify the benefits, and the sources of the benefits in particular, of sanitation infrastructure, as well as the sources of these benefits to develop more efficient and greater policies and with higher impact.

We estimate a model for the individual production function of health for children in rural India assuming a linear-in-parameters approximation. The main inputs to the function are the household's access to sanitation and the ratio of access to sanitation (i.e., coverage) at the village level. This model is estimated using a sample of 206,414 children from households in rural villages of India from the District Level Household Survey (DLHS-3) 2007–08. One of the novel features of this model is that we have a measure for the estimation of sanitation externalities' impact through use of a robust proxy for village access to sanitation facilities. Using this approach, we can separate the direct from the externality benefits of sanitation external ones, and analyze how households benefit from different levels of sanitation coverage.

This paper contributes to the literature of the health impact of sanitation access in two ways. First, it adds to the existing evidence on the positive relationship between better sanitation infrastructure and early childhood health outcomes (Bose 2009; Gunther & Fink 2010; Kumar & Vollmer 2012; Patil *et al.* 2014; Duflo *et al.* 2015; Geruso & Spears 2015). The second, and most important, adds evidence to the idea on the existence of that positive externalities derive from access to sanitation infrastructure.

Our principal empirical findings are as follows:

1. In a sample where the average prevalence of diarrhea among children under 48 months is 12.1%, we find a reduction of 10% (or 1.26 percentage points (pp) in probability) of diarrhea explained by gaining access to improved sanitation. We also estimate a reduction in probability of diarrhea on the last 2 weeks of 5% (or 0.63 pp) when moving from open defecation to fixed-point defecation.
2. We find evidence that supports a concave relationship between sanitation coverage at village level and diarrhea prevalence, whereby the positive externalities are only enjoyed after a certain level of sanitation coverage has been achieved. The potential benefits are greater at higher levels of coverage due to the reduction in the probability of contact with human excreta. The estimates indicate that there is no improvement at all until the 30% coverage is achieved and that half of the potential total gains are only reached when coverage is approximately 75%.
3. In terms of the magnitude of the effect from both sources of benefit, direct and external ones, we estimate that a child who moves from a household without improved sanitation and a low ratio of village access to a household with improved sanitation and a high ratio of village access enjoys a reduction in diarrhea prevalence of 47% (from 12.54% to 6.79%). For the case of fixed-point defecation, the overall effect is a reduction of 26% (from 12.49% to 9.25%).
4. We find the external benefit is largely responsible for the health improvements when we separate the contribution of the direct and external benefit in overall reductions of diarrhea prevalence. While for improved sanitation the direct benefits account for 23%, for fixed-point defecation the direct benefits accounts for just 19% of the overall gains.
5. We find no significant difference (but negative effect) in terms of health outcomes between children living in households whose members practice open defecation and those who are living in households with unimproved sanitation, which includes unimproved sanitation facilities and those who are sharing either improved or unimproved facilities with other households. The gains are only for children who live in households with access to improved sanitation as defined by the World Health Organization/United Nations Children's Fund (WHO/UNICEF) Joint Monitoring Programme (JMP) for Water Supply and Sanitation.

We subject the results from our parametric model to several robustness checks to test whether characteristics besides access to sanitation, and which vary systematically between households, could be driving our results leading some households to invest in sanitation at higher rates while simultaneously taking precautions to lower diarrhea

risk in children. We find that our results hold under each of these alternative econometric specifications.

Specifically, we use quasi-experimental propensity score matching to first compare children living in low-ratio of access villages with children living in high-ratio of access villages, and second to compare children with and without access to sanitation at different ratios of village access. In a third test, we estimate our baseline parametric model using an outcome, prevalence of cough, which we hypothesize could be influenced by unobserved characteristics such as income, healthier parenting behavior, or risk tolerance in a similar way as diarrhea. To further test whether an unobservable factor such as income could be driving our results, we substitute access to electricity, an alternative household infrastructure asset, for sanitation. Finally, we estimate a full counterfactual model, which compares children with and without access to sanitation in villages with no coverage to children living in households with and without access in villages at 10% increments of rate of access to sanitation.

The remaining sections of this paper are organized as follows: first, the paper presents the motivation for trying to understand the source of benefits of sanitation, which will then be complemented with a review of the literature on household sanitation and its relation to children's health and welfare. Next, a theoretical model to understand the direct and external benefits of sanitation will be described, followed by the empirical implementation and the description of the data used for this paper. Finally, the empirical results of the paper will be presented, including several econometric specifications used to test the robustness of the results. The paper ends with a discussion of the implications of these results for sanitation policy.

MOTIVATION

One of the principal motivations underlying this paper is the fact that the Millennium Development Goal (MDG) target to halve the proportion of the population that lacked access to basic sanitation was not met in 2015. Globally, there are still 2.4 billion people without access to sanitation and India alone is home to over 30% of this population. There is little debate that safe sanitation improves health. However, despite a theoretical model that demonstrates that at least

a proportion of these health improvements stems from community benefits to improved sanitation, household sanitation is still often perceived as only a private good. The consequence of this is that many governments have adopted policies that shift the burden of investment to households. The resulting underinvestment by households suggests they either have credit or other financial constraints that prevent investment in improved sanitation or they underestimate the private benefit to sanitation. (We do not dismiss that it could be problems from the supply side which does not meet the needs of consumers.)

The second motivation for studying the benefits of household sanitation has to do with the ubiquity of policies aimed to generate open defecation-free (ODF) communities. These policies are based on the plausible assumption that so long as some human excreta remains in the open, communities will continue to be at risk of fecal-borne diseases, such as diarrhea. This assumption recognizes that open defecation poses a negative externality on others. One of the most emblematic programs calling for ODF communities is the Total Sanitation Campaign (TSC) in India (see Appendix A for details, available with the online version of this paper), the site of this study.

'Elimination of open defecation' is now a post-2015 Sustainable Development Goal, even when there is little evidence which shows that a community that has achieved 100% ODF is better off from a health perspective. Surprisingly, there is still no quantitative evidence that the only way to enjoy full health benefits from improved sanitation is having full coverage at the community level. Therefore, quantification of the relationship between access to sanitation both at the household and community levels, and the role of externalities as a channel for benefits can inform a debate around whether public investment or subsidies are justified for what is normally perceived as a private good. If externalities are an important channel through which households benefit, then private markets alone will fail to produce the optimal social equilibrium that maximizes the return of sanitation investments in society.

LITERATURE REVIEW

It is estimated that the risk of fecal-oral diseases resulting from poor water, sanitation and hygiene (WSH) contributes

to 5.7% of the global disease burden (Prüss *et al.* 2002). Diarrhea is the most common fecal-oral disease associated with poor WSH, accounting for nearly 0.8 million child (under five years) deaths each year worldwide (Liu *et al.* 2012), and 11% of all deaths of children under five in 2010 (UNICEF 2012). Open defecation, practiced by an estimated 626 million people in India (JMP 2012) and around 560 million by 2015, is indicated as a major contributing factor to diarrheal disease. Evidence therefore suggests that stopping open defecation and safely containing feces with improved household sanitation can effectively break down the fecal-oral transmission of disease, thereby improving health. Many studies have attempted to quantify the effect of this transition from widespread open defecation to improved sanitation on health outcomes. However, to date, few rigorous studies have been conducted and most have taken place alongside complementary water and hygiene interventions, making it difficult to isolate the effect of safe feces containment (Andrés *et al.* 2013). Indeed, some have suggested there is too little evidence on the complementary effects of WSH interventions and that too much attention has been focused on understanding the epidemiology of single transmission pathways (see Eisenberg *et al.* 2007, 2012). A widely cited meta-analysis estimates the impact of sanitation on diarrheal disease to be 32% (Fewtrell *et al.* 2005), but this estimate is based on just two studies.

More recently, communities which were randomly assigned to receive a community-led total sanitation and sanitation marketing intervention reported reductions in 2-day prevalence of diarrhea in children under five of 45% (Cameron & Shah 2013). It is critically recognized that the consequences of poor sanitation extend beyond the burden of diarrhea. Exposure to fecal contamination in the environment and chronic diarrhea decreases the ability to absorb essential nutrients, leading to malnutrition (Checkley *et al.* 2008). This effect also goes the other way, with poor nutritional status associated with increased risk of diarrhea (see for example, Chen *et al.* 1981). Poor sanitation is increasingly recognized as a distal cause of underweight and stunting in children (Humphrey 2009). Malnourishment during childhood is associated with severe long-term consequences. These long-term consequences include poor cognitive development, lower school attendance, reduced human capital

attainment, and potentially, a higher risk of chronic disease in adulthood (Victora *et al.* 2008).

There is strong evidence that the individual benefit to having improved sanitation is potentially very large, but to what extreme is this benefit actually driven through a community benefit, or externality to private sanitation, is less well understood. It has been suggested that even a few remaining open defecators in a community risk bringing fecal matter back into the environment resulting in continued contamination of soil and water, and potentially infecting others in the community who practice safe sanitation. In other words, even if my household has chosen to practice safe sanitation, if others in my community continue to defecate in the open, what are the consequences for the welfare of my household?

Experimental and quasi-experimental studies have looked at the impact of improved sanitation on child health outcomes including diarrhea, infant mortality, and stunting. However, most of these studies have focused on the effects at the household level, and while they acknowledge the possibility that an important part of the effect has been driven by externalities, previous work has not estimated these effects (see for example, Spears 2013; Spears *et al.* 2013). Recent research such as that of Patil *et al.* (2014) and Clasen *et al.* (2014) may have even underestimated the beneficial effects of improved sanitation by restricting the analysis to effects over the average at village level, which mixes the effects from externalities and private returns to sanitation. Headey (2013) investigates the drivers of rapid declines in child undernutrition in Bangladesh, and finds evidence of a positive and non-linear relationship between levels of open defecation in a community and nutrition outcomes.

One of the most relevant analyses for the current study is a recent paper that used propensity score matching to estimate the effect of access to improved sanitation in India on diarrhea in children under five using the DLHS-3 (Kumar & Vollmer 2012). The authors found that access to improved sanitation reduces diarrhea by 2.2 pp corresponding to a 17% drop in prevalence. Another study (Bose 2009), using similar methods from Nepal, looks at access to sanitation using a 2006 Demographic and Health Survey (DHS), finding reductions of 5% from mean prevalence of 8%, a substantially larger relative reduction. Similar results were

found in Duflo et al. (2015), when using experimental data to evaluate a water and sanitation program in India, reductions between 30% and 50% in diarrhea prevalence were found.

In other studies, Spears (2013) estimates the effect of open defecation in anthropometric measures between and within countries, and finds positive benefits from externalities and direct effects of eliminating open defecation; also, Spears (2012) looked at the relation between infant mortality and sanitation and found that an important driver of the effects was coming through externalities. Finally, Gunther & Fink (2010) analyzed data from 172 DHS surveys and found that households having flush toilets have 13% lower odds of diarrhea than households without – in line with the estimates from Kumar & Vollmer (2012), but much less than estimates from meta-analyses. Moreover, this study finds the benefits at the cluster level (a cluster in the DHS is typically a village for rural areas and district for urban areas) are twice as large as those at the household level, suggesting there may be positive externalities to improved sanitation. In contrast to the positive results found in these studies, a study in India that employed multiple matching techniques failed to find a consistent effect of access to improved toilets on diarrhea (Fan & Mahal 2011). The authors hypothesize that behavior, rather than infrastructure, could be driving these null results.

MODEL

In this paper we seek to answer two questions: (i) what are the private benefits of a household's access to sanitation infrastructure on early childhood health outcomes? and (ii) what are the benefits produced through an external effect, or externalities, of a neighbor's access to this sanitation infrastructure? For this we start from the following health production function presented in Equation (1):

$$y_{ijk} = f(s_{ijk}, e(s_{-ijk}), x_{ijk}) \quad (1)$$

The health production function of the child y_{ijk} depends on the sanitation investments s_{ijk} of the household, an environmental factor or externality e produced as byproduct of sanitation investments on other households, and a vector

of characteristics of the child and the parents. In this paper we understand the sanitation externality $e(s_{-ijk})$ as the one indicated by Gersovitz & Hammer (2004): where an infectious person, due to his own decision, infects other persons (pure infection externality); or the preventive action of investing in sanitation infrastructure reduces the probability of someone else being infected unconditional of their decisions of investment in sanitation infrastructure (pure prevention externality).

From Equation (1), the parameters of interest to answer the previously proposed questions, are:

1. the difference between children with access to certain sanitation infrastructure

$$E[y_{ijk}|s_{ijk} = 1] - E[y_{ijk}|s_{ijk} = 0] \quad (2)$$

2. the size and rate of changes in levels of externality

$$\frac{\partial y_{ijk}}{\partial e(s_{-ijk})} \quad (3)$$

$$\frac{\partial^2 y_{ijk}}{\partial e(s_{-ijk})^2} \quad (4)$$

Empirical implementation

We test the hypothesis presented in Equations (2) and (3) for different types (in terms of quality) of access to sanitation infrastructure, namely access to improved sanitation facilities, access to fixed-point defecation, and open defecation, based on the theory that as quality of sanitation improves so does health. These levels are defined from the 'Sanitation Ladder' (Figure 1). The rungs on the ladder represent an improvement in sanitation infrastructure, which correlates with gains in health outcomes. We define access to sanitation according to the definitions used by the JMP (World Health Organization/United Nations Children's Fund Joint Monitoring Programme for Water Supply and Sanitation website, <http://www.wssinfo.org>), which tracks country progress towards achieving the MDGs related to water and sanitation. The JMP uses two key characteristics to classify access to sanitation: (a) degree to which infrastructure

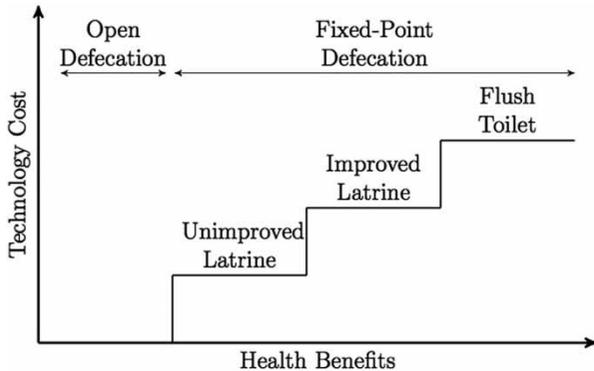


Figure 1 | The Sanitation Ladder.

safely isolates excreta from the environment and (b) ownership of the structure. Table 1 shows how each type of sanitation is classified following JMP's criteria.

In terms of isolation of excreta from the environment there are three levels. The lowest level is open defecation in the bush, fields, or surface water. The second includes unimproved facilities, such as a pit latrine without a slab, a hanging toilet or latrine over a pond or other surface water, or a flush toilet that does not properly dispose of waste. These types of sanitation do not hygienically separate human excreta from the environment. Finally, improved facilities, including septic or piped sewage, pit latrines with slab, and composting toilets, hygienically separate human

excreta from the environment. In terms of ownership, the JMP classifies any shared latrine as unimproved sanitation due to the constraints placed on households having to share facilities. From the improved facilities' definition, and using the ownership structure, we can define four groups that do not practice open defecation: (a) shared unimproved facilities, (b) private unimproved facilities, (c) improved shared facilities, and (d) improved sanitation.

We define a child's access to sanitation S_{jk}^v as a dichotomous variable in Equation (5), which takes value 1 if the child's household j from village k has access to sanitation, and zero if otherwise. We define this for two types of access to sanitation $v = [1$ (Fixed-Point Defecation), 2 (Improved Sanitation)] according to the JMP's classification. (The actual question on the survey is phrased as 'What kind of toilet facility do members of your household usually use?' The distinction on the word 'use' alleviates in a way the problems on the difference between property and usage (Gupta et al. 2014).)

$$S_{jk}^v = \begin{cases} 1 & \text{if household } j \text{ of village } k \text{ has access to sanitation } v \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

Using our child's household access to sanitation variable S_{jk}^v , we define R_k^v in Equation (6) as the proportion of

Table 1 | Classification of sanitation infrastructure access

Open defecation No facilities, bush or field	Fixed-point defecation	
	Improved sanitation	Unimproved sanitation
	Private flush toilet	Shared flush toilet
	Private piped sewer system	Shared piped sewer system
	Private septic tank	Shared septic tank
	Private flush/pour flush to pit latrine	Shared flush/pour flush to pit latrine
	Private ventilated improved pit latrine (VIP)	Shared ventilated improved pit latrine (VIP)
	Private pit latrine with slab	Shared pit latrine with slab
	Private composting toilet	Shared composting toilet
		Flush/pour flush to elsewhere, private or shared
		Pit latrine without slab, private or shared
		Hanging toilet or hanging latrine, private or shared

Note: This classification is based on the definitions created by the WHO/UNICEF JMP for Water Supply and Sanitation for the MDG monitoring. An improved sanitation infrastructure is one that hygienically separates human excreta from human contact and is for private use of the household.

households in the village j that have access to the sanitation infrastructure v . So, variable R_k^v measures the externality through the level of decontamination at the village level. The distribution of the ratio in our sample is presented in Figure 2.

$$R_k^v = \frac{\sum_j S_{jk}^v}{J} \tag{6}$$

In order to estimate the parameters of interest we use a parametric model assuming a linear-in-parameters approximation of Equation (1). We impose a quadratic linear relation between our measure of externality and our outcome of interest based on an exploratory data analysis using a local polynomial regression estimator (Nadaraya 1964; Watson 1964) between ratio of access to sanitation and our health outcome measure, diarrhea prevalence.

Then, we evidence a nonlinear relationship between the two variables (Figure 3(a) and 3(b)). After which, we estimate the parametric model presented in Equation (7) with a probit model:

$$D_{ijk} = \alpha + \theta S_{jk}^v + \delta^{00} R_k^v + \delta^{01} (R_k^v)^2 + \delta^{10} (R_k^v) * S_{jk}^v + \delta^{11} (R_k^v)^2 * S_{jk}^v + \beta X_{ijk} + \epsilon_{ijk} \text{ for } v = 1, 2 \tag{7}$$

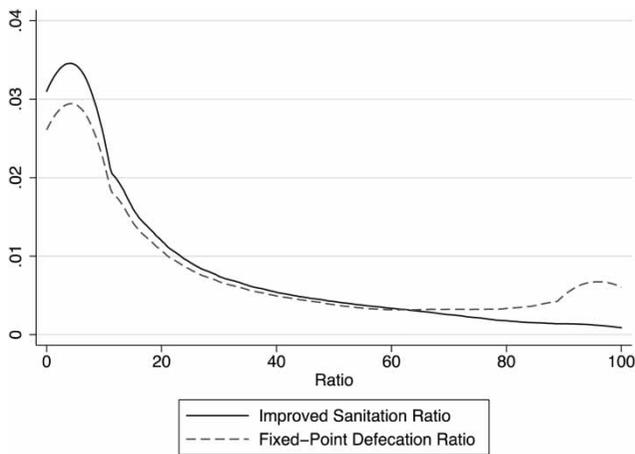


Figure 2 | Distributions of improved sanitation ratio. Note: The ratios of Improved Sanitation and Fixed-Point Defecation are constructed as the proportion of households in the PSU, defined as village in DLHS-3, that has access to that kind of sanitation technology. The use of villages as the PSU for rural areas by DLHS gives us a random sample of households of each village, making the ratio reliable. The distribution presented in the figure was estimated using a kernel density estimation with a bandwidth of 5.

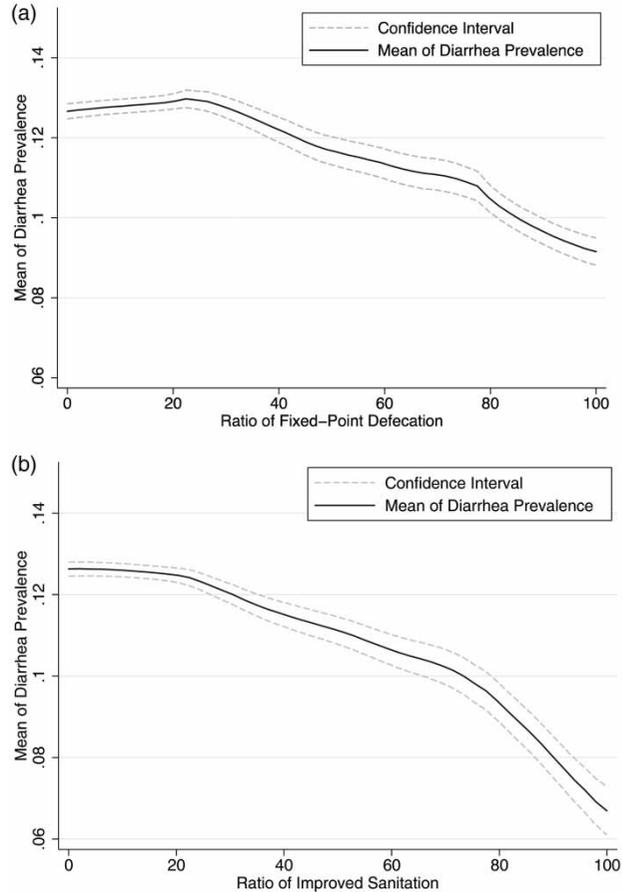


Figure 3 | Nonparametric relation of diarrhea and ratio of access to sanitation. (a) Ratio of fixed-point defecation. (b) Ratio of improved sanitation. Note: The curve presents the local polynomial regression between the prevalence of diarrhea for each level of ratio of access to sanitation. The pointed lines are the confidence interval at 95%.

We assume different benefits from externalities to children that live in households with and without access to sanitation infrastructure. We include X_{ijk} , a vector of controls including years of schooling, sex and caste of household head, age in months of the child, and gender of the child. Because we use a variable that is measured on a higher level than the dependent variable, we will estimate our model using clustered errors (for further discussion, see Moulton (1990) and Bertrand et al. (2004)) at the village level to obtain robust errors that do not affect our inference.

Using this model to test the hypothesis presented in Equations (2) and (3), our relevant parameters will be: the marginal effects on the probability of diarrhea of access to fixed-point defecation (S_{jk}^1) and of access to improved sanitation (S_{jk}^2). The first order derivative of the ratio of access over the

probability of diarrhea gives us the change in diarrhea benefit from externalities, and the second order derivative of the ratio of access over the probability of diarrhea gives us the rate of change of the benefit from externalities.

These parameters allow us to simulate the predicted probability of diarrhea for the different combinations of ratio of access and household access to sanitation technology.

DATA

We use a subsample of 209,762 children between 0 to 48 months of age that live in rural areas from the Third Round of the DLHS-3. The DLHS is a nationwide survey with district level representation of India's households, which collects information on family planning, maternal and child health, reproductive health of ever married women and adolescent girls, and utilization of maternal and child healthcare services.

As described previously, an important feature of this paper is the explicit identification of the external benefit of sanitation, or externality, as measured through the ratio of access to sanitation at the village level. In this paper, we exploit the sample design used in DLHS-3, where the primary sampling units (PSU) for the rural areas are villages (International Institute for Population Sciences 2010). A household census was conducted in each PSU, which gives us the actual rate of access to the different sanitation technologies at the village level. However, because PSU's for urban areas are not clearly delineated (that is, neighbors are not clearly defined), we restrict the sample to rural areas only. In our sample, the ratio of access to sanitation ranges from 0 to 100%. Our main health outcome measure is a binary variable (Equation (8)) that takes value 1 if a mother reports that her child 'has had diarrhea in the last two weeks', and 0 if otherwise.

$$D_{ijk} \begin{cases} 1 & \text{if has had diarrhea in the last two weeks} \\ 0 & \text{if has not had diarrhea on the last two weeks} \end{cases} \quad (8)$$

When we analyze the differences between those who have access to improved sanitation and fixed-point

defecation, the differences in diarrhea prevalence emerge. While the limitations of caregiver-reported diarrhea have been noted in the literature (Das et al. 2012), we chose this outcome in order to make the findings more comparable to previous studies on the health benefits of sanitation.

Figure 4 shows the distribution of type of sanitation for this sample. We find that 18% of children under 48 months live in households with improved sanitation, 10% have unimproved facilities which include shared improved facilities, and 72% live in households which lack access to any type of sanitation facility and thus defecate in the open. These ratios are reasonably comparable with the estimations of JMP for rural areas in India for 2007. Table 2 presents the summary statistics of these variables for the 209,762 observations in our sample, cross-tabulated by type of access to sanitation. In our sample we observe that the average prevalence of diarrhea is 12.1%.

EMPIRICAL ESTIMATES

In this section, we discuss our empirical findings: first for direct benefits on childhood prevalence of diarrhea resulting from access to sanitation, and second the benefits resulting from externalities.

Direct benefits

Table A1 (available with the online version of this paper) presents the marginal effect on diarrhea prevalence of moving from open defecation to fixed-point defecation (S_{jk}^1), and of moving from unimproved sanitation (which includes open defecation) to improved sanitation (S_{jk}^2), estimated from model 4 in Tables 3 and 4 which correspond to Equation (8). Similar to other studies, we find a direct benefit to children living in both households that stopped defecating in the open and moved to fixed-point defecation or as well as for those who moved to improved sanitation.

The magnitude of the effect is a 0.62 pp reduction in diarrhea prevalence for a child in a household that goes from open defecation to fixed-point defecation (unimproved or improved facilities). For the same children, a direct move from unimproved sanitation or open defecation to a sanitation facility classified as improved sanitation provides a

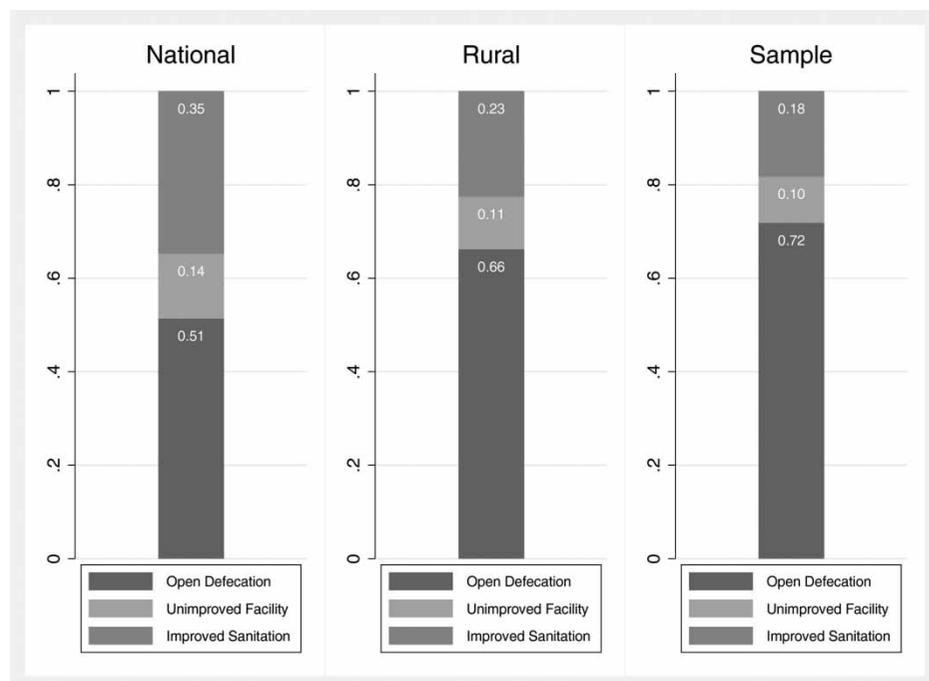


Figure 4 | Distributions of sanitation infrastructure. *Note:* Open defecation is defined as if the household does not have any access to any type of facility. The unimproved facilities are sanitation infrastructures that do not separate human excreta from human contact and/or are shared with other households. Improved sanitation are sanitation infrastructures that separate human excreta from human contact and are private for the household. These definitions come from the WHO/UNICEF JMP for Water Supply and Sanitation for the MDG monitoring. This figure is constructed with the sample of children used for the estimations.

Table 2 | Summary statistics

Variable	Full sample		With access to improved sanitation		With access to fixed-point defecation	
	Mean	(Std. dev.)	Mean	(Std. dev.)	Mean	(Std. dev.)
Diarrhea	0.121	(0.327)	0.102	(0.303)	0.106	(0.308)
Years of schooling of household head	4.452	(4.525)	6.483	(4.805)	6.116	(4.737)
Access to fixed-point defecation	0.282	(0.450)	1.000	(0.000)	1.000	(0.000)
Access to improved sanitation	0.182	(0.386)	1.000	(0.000)	0.645	(0.478)
Ratio of access to fixed-point defecation	0.036	(0.186)	0.000	(0.000)	0.127	(0.333)
Ratio of access to improved sanitation	28.39	(33.07)	60.95	(32.64)	65.66	(32.24)
Gender of household head (male = 1)	18.38	(23.31)	45.66	(28.36)	39.60	(27.74)
Household head is from a low caste	0.084	(0.277)	0.092	(0.289)	0.084	(0.278)
Access to improved facilities are shared	0.191	(0.393)	0.133	(0.339)	0.122	(0.327)
Age of the child (in months)	22.79	(13.77)	23.04	(13.72)	23.19	(13.71)
Gender of child (male = 1)	0.521	(0.500)	0.528	(0.499)	0.526	(0.499)
Number of observations	209,762		37,275		57,311	

Note: Children with missing values for the variables presented are excluded from the sample. The variables gender of the household head and gender of the child takes value 1 when it is a male and zero when it is female.

Table 3 | Fixed-point defecation estimations

	Model 1	Model 2	Model 3	Model 4
Access to fixed-point defecation	-0.107***	-0.0967***	-0.0289	-0.0195
Sex of household head		-0.0161		-0.017
Years of schooling of household head		0.000376		-0.00013
Household head is from a low caste		0.0549***		0.0509***
Age of the child (in months)		-0.0109***		-0.0109***
Gender of child (male = 1)		0.026***		0.0259***
Ratio of fixed point-defecation			0.00235**	0.00235**
Ratio of fixed point-defecation x access			0.00114	0.000847
Square ratio of fixed point-defecation			-0.0000406***	-0.0000398***
Square ratio of fixed point-defecation x access			-0.0000271**	-0.0000213*
Constant	-1.14***	-0.932***	-1.15***	-0.94***
No. of observations	209,762	209,758	209,696	209,692

Note: The table presents the coefficients from the estimation of the parametric model with clustered errors at the PSU level. Significance levels: *10%, **5%, and ***1%.

Table 4 | Improved sanitation estimations

	Model 1	Model 2	Model 3	Model 4
Have access to improved sanitation	-0.12***	-0.115***	-0.0679*	-0.0622*
Gender of household head (male = 1)		-0.0149		-0.00945
Years of schooling of household head		0.000271		0.000265
Household head is from a low caste		0.059***		0.0597***
Age of the child (in months)		-0.0109***		-0.0109***
Gender of child (male = 1)		0.026***		0.0259***
Ratio of improved sanitation			0.00126	0.00139
Ratio of improved sanitation x access			0.00145	0.00138
Square ratio of improved sanitation			-0.0000431**	-0.0000414**
Square ratio of improved sanitation x access			-0.0000431***	-0.0000417**
Constant	-1.15***	-0.938***	-1.15***	-0.94***
No. of observations	209,762	209,758	209,696	209,692

Note: The table presents the coefficients from the estimation of the parametric model with clustered errors at the PSU level. Significance levels: *10%, **5%, and ***1%.

greater benefit of 1.26 pp reduction in diarrhea prevalence. Given that diarrhea prevalence in the sample is 12.1%, this implies relative reductions of between 5.1% and 10.4% for fixed-point defecation and improved sanitation, respectively. We should analyze these results carefully because access to fixed-point defecation and access to improved sanitation are not mutually exclusive: improved sanitation is a subset of fixed-point defecation.

Does the benefit from fixed-point defecation stem from the isolation of excreta itself (type of sanitation infrastructure), the ownership (sanitation facility is shared with other households or used only by household members) or a combination of these (improved sanitation)? To test the contribution of these factors for the direct benefits on diarrhea, we run our parametric model for the marginal effect on diarrhea prevalence of each ownership and

infrastructure type against open defecation. We use S_{jk}^* , defined in Equation (9), and control for household coverage of each type. The results are shown in Figure 5.

$$S_{jk}^* \begin{cases} 0 & \text{if household } j \text{ of village } k \text{ does not have} \\ & \text{access to any facility} \\ 1 & \text{if household } j \text{ of village } k \text{ has} \\ & \text{access to shared unimproved sanitation} \\ 2 & \text{if household } j \text{ of village } k \text{ has access to private} \\ & \text{unimproved sanitation} \\ 3 & \text{if household } j \text{ of village } k \text{ has access to} \\ & \text{shared improved sanitation} \\ 4 & \text{if household } j \text{ of village } k \text{ has access to} \\ & \text{improved sanitation} \end{cases} \quad (9)$$

The graph illustrates that there is no significant benefit (or detriment) in terms of reduction of diarrhea for any unimproved sanitation, as classified by the JMP. In other words, in a context where externalities are included as a source of benefit, the categories of unimproved/private, unimproved/shared, and improved/shared are no better than open defecation. The DLHS-3 data do not provide details on the number of households that share the latrines reported, whether the shared latrine belongs to the household or another household. These details could help us

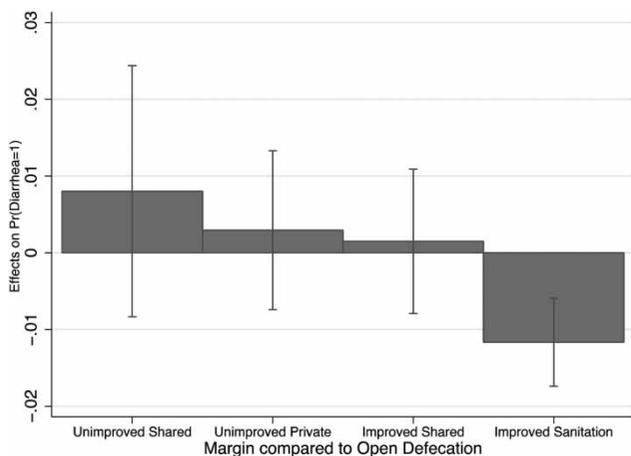


Figure 5 | $(\partial Pr_{jk}(D = 1))/(\partial S_{jk}^*)$. Note: The bars presented are the difference between the option t and option 0 of sanitation, as following $Pr(D_{ijk} | S_{jk}^* = t) - Pr(D_{ijk} | S_{jk}^* = 0)$ where $S_{jk}^* = 0$ is open defecation. The categories presented are defined by the combination of two conditions, isolation of excreta from human contact and if it is for private use for the household, where improved sanitation is the JMP classification which achieves both conditions. The confidence intervals are constructed using a 95% significance.

understand the intensive margin of the effect. This evidence is relevant for the TSC, where the goal is eliminating open defecation, and where it is evident that the particular fixed-point defecation offered matters. If the way to achieve the goal is building a fixed-open defecation infrastructure different from improved sanitation, this investment would have a negative rate of return because it is not producing any visible impact.

Externalities

Taking into consideration the effect of externalities estimated through Equation (8), we find there is a positive and significant relationship between the level of access to sanitation in a village and the prevalence of diarrhea (Table A1). Moreover, we corroborate our hypothesis that this relation follows a concave function, where at low levels of village access there are limited positive externalities, but as access to sanitation increases, these externalities begin to accrue at a faster rate. Again, the findings are similar for fixed-point defecation and for improved sanitation. An interesting finding is the difference between the children that already have access and the ones who do not. We observe that the change in diarrhea prevalence caused by the external effect is larger for children living in households without access, but the rate of change of this effect is larger on children living in households with access.

In Figure 6(a) and 6(b) we illustrate the predicted probability of diarrhea in the past 2 weeks from Equations (13)–(17) conditional on a household’s private access to sanitation and the ratio of access to sanitation at the village level, for both fixed-point defecation and improved sanitation, respectively. In both cases, we confirm our hypothesis that there is a negative relationship between the ratio of access and diarrhea, and more importantly, that the benefit from externalities to sanitation follows a concave function. The rationale behind these results is that a sufficient number of households must have access to improved sanitation in order to decontaminate the village to a level where everyone benefits. This convergence in probability shows that children living in households without access to sanitation enjoy nearly the same benefit from lower levels of open defecation as those who have sanitation facilities.

This finding bolsters the policy argument for high coverage as a goal. The benefits that a child would receive when he goes from a state of no sanitation coverage ($S_{jk}^2 = 0$) in a village where no one has access to improved sanitation ($R_k^2 = 0$) to a state of access to improved sanitation ($S_{jk}^2 = 1$) and full coverage at the village level ($R_k^2 = 100$) is estimated from our simulations to yield a decrease in the prevalence of diarrhea of 5.7 pp, which is a 47.5% reduction in the prevalence of diarrhea. This final result shows that 23% of the combined health benefits of improved sanitation are explained by the direct benefit (i.e., vertical distance between the two curves in Figure 6(b)) and 77% are explained by the indirect or externality benefit (i.e., horizontal distance between 0% and 100% ratio of access

in Figure 6(b)). The estimates indicate that there is no improvement at all until 30% of improved sanitation coverage is achieved and that half the potential total gains are only reached when coverage is 75%. This finding could provide one candidate explanation for the lack of health impacts found in Patil et al. (2014), Clasen et al. (2014), and Hunter & Prüss-Ustün (2016) since none of these intervention studies led to coverage levels above 65%.

Taking into consideration the fact that our estimates show a smaller direct impact of sanitation on child health than found in previous research, we conclude that direct benefits have previously been overestimated because they include a portion of the external benefit. In addition, by not accounting for externalities, previous research may have underestimated the overall benefit.

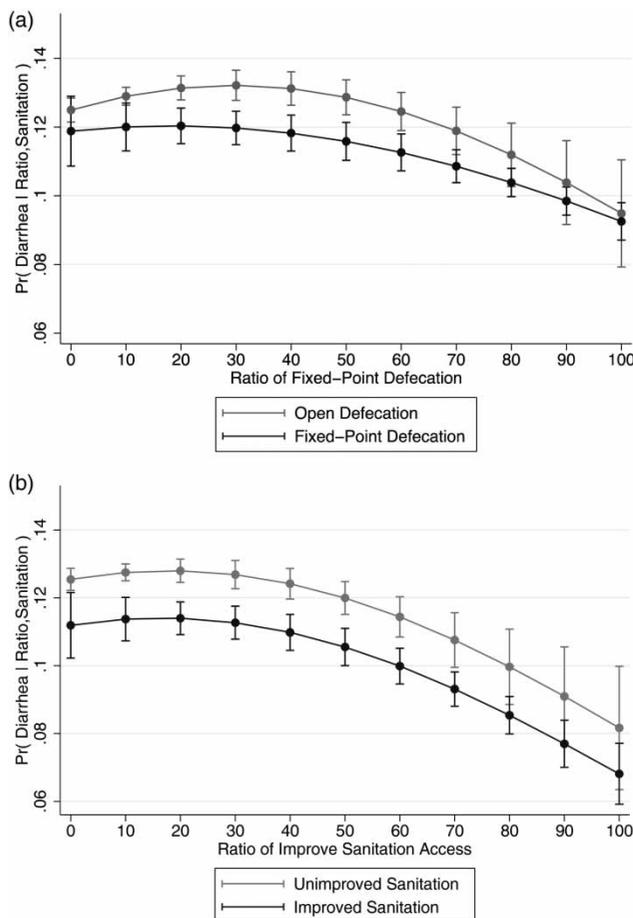


Figure 6 | Parametric model simulation. (a) Fixed-point defecation. (b) Improved sanitation. Note: Calculated as $Pr(D_{ijk} | R_k^2, S_{jk}^2) | R_k^2 = J$ for each value of improvement sanitation ratio, $J \in [0, 100]$. The bars at each point of estimation correspond to the confidence interval using robust clustered errors at the village level.

ROBUSTNESS CHECKS

In our analysis we control for a set of household and child characteristics that are correlated with our main outcome of diarrhea prevalence and the household's probability of having access to sanitation. However, we are unable to control for characteristics that are unobservable, like risk aversion, that vary systematically between households, leading some households to invest in sanitation at higher rates while simultaneously taking precautions to lower diarrhea risk in children. These unobserved characteristics could bias our results. To test the robustness of the results presented in the previous section, we run four separate checks: (i) difference in health outcomes between children in households at either extreme of the distribution of ratio of access; (ii) difference between households with and without access to sanitation at similar levels of village ratio of access; (iii) estimation of our model using a placebo treatment and outcome as falsification test; and finally (iv) a counterfactual model.

Testing effects at the extremes of ratio of access

The first exercise to check the robustness of our results tests whether differences besides access to sanitation, such as income levels, might explain the difference in health outcomes between children living in villages in which almost everyone

openly defecates with children living in villages that are almost or completely ODF. If children across such villages are not comparable, then the difference in diarrhea that we observe would be biased and not sufficiently explained by access to sanitation. Assuming non-random selection in the level of access to sanitation at the village level, we estimate the treatment parameters $E(D(R_k^* = 1) - D(R_k^* = 0))$ (average treatment effect, ATE), $E(D(R_k^* = 1) - D(R_k^* = 0) | R_k^* = 1)$ (average treatment on the treated, ATT), and $E(D(R_k^* = 1) - D(R_k^* = 0) | R_k^* = 0)$ (average treatment on the untreated, ATU) using a nearest neighbors matching estimator (Leuven & Sianesi 2012) between children who live in villages with a ratio of access lower than 10% and children who live in villages with a ratio of access higher than 90%, defined in Equation (10). Comparing the extremes of the distribution of access to sanitation we test the robustness of the slope on Figure 6(a) and 6(b).

$$R_k^* = \begin{cases} 0 & \text{if } R_k^* < 10\% \\ 1 & \text{if } R_k^* > 90\% \end{cases} \quad (10)$$

We match 15 control observations for each treatment observation for both open defecation and improved sanitation, using the same controls as in our parametric model (X_{ijk}), which includes a wealth index. Figure 7 shows the three

treatment parameter estimates. Through matching, we estimate that children living in villages where the ratio of access to fixed-point defecation is greater than 90% have a 2.65 lower prevalence of diarrhea (ATE) than children living in villages where the ratio of access to fixed-point defecation is less than 10%. Diarrhea prevalence is similarly lower (ATE 6.32 pp) for children living in villages where the ratio of access to fixed-point defecation is greater than 90%. Both the magnitude and direction of the results are consistent with the differences presented in our parametric model when moving from a state of open defecation ($R_k^v \approx 0$) in a village where everyone practices open defecation ($R_k^v \approx 100$) to a state of fixed-point defecation/improved sanitation in an ODF village.

A second, more policy-oriented finding from this exercise is the difference in the effect we observe for those at the top of the distribution of access (ATT) and the effect for those at the bottom of the distribution (ATU). In a context of complete randomness of the treatment we could expect these treatment parameters would be the same, however we observe that the benefits are largely captured by those who are in a treated state. Nevertheless, there are still benefits to be captured by moving a child who lives in a village with very low access to one with high access. This suggests that the ATE is driven both by the potential

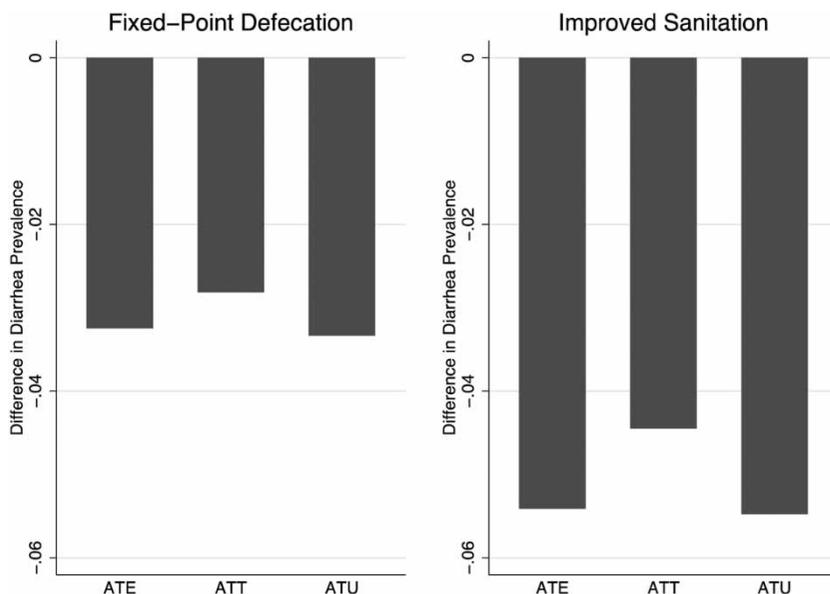


Figure 7 | Matching estimators between extremes of ratio of access. Note: The estimators presented are estimated using a nearest neighborhood matching estimator, with a set of 15 matches for each observation in the treatment group. The six estimators are significant at 95% of confidence. The propensity score distributions are presented in Figure A1 in Appendix C (available with the online version of this paper).

gains for those without access as well as the gains enjoyed by those who already have access.

Testing effects within ratio of access

At each ratio of access in a village we only observe health outcomes for households in the village with and without access to sanitation and cannot observe health outcomes for the same household residing in a village with a different ratio. To generate a counterfactual for the effect of household access to sanitation within a given ratio of access we implement a nearest neighborhood matching estimator between households at each village ratio of access (Equation (11)) where $R_k^{v*} \in r^* = [10, 80]$.

$$\begin{aligned} \frac{\partial D_{ijk}}{\partial S_{jk}^v} \Big|_{R_k^{v*} = (r^*, r^* + 10)} &= E [D_{ijk} | S_{jk}^v = 1, R_k^{v*}] \\ &= (r^*, r^* + 10) - E [D_{ijk} | S_{jk}^v = 0, R_k^{v*} = (r^*, r^* + 10)] \end{aligned} \quad (11)$$

The graph in Figure 8 demonstrates that holding the ratio of access constant, the within-direct benefit of moving from no sanitation to improved sanitation (or fixed-point sanitation) is similar to the direct benefit that is predicted in our parametric model and shown in Figure 6(a) and 6(b). Our results indicate that as the ratio of access at the village level increases, there are fewer gains to be made by a household moving from open defecation to improved sanitation (fixed-point defecation), highlighting the role of externalities. Additionally, we

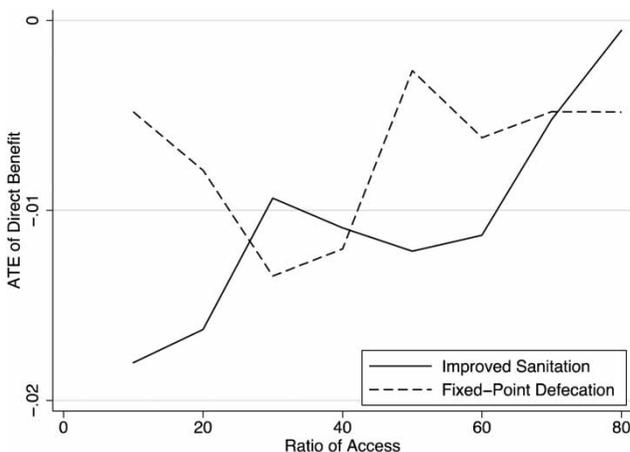


Figure 8 | Direct benefit matching within level of access.

find large direct benefits to moving from open defecation to fixed-point defecation when the ratio of access at the village level is between 30 and 40%.

Falsification tests

The third potential problem we encounter stems from the fact that our measure of the external benefits to improved sanitation could be explained by some unobservable characteristic that is correlated with both access to sanitation and assets for a healthier environment or healthier parenting behavior. If this is the case, the difference in diarrhea prevalence estimated in our parametric model would be partially or fully explained by differences in this unobservable characteristic. To test the hypothesis that some other factor, such as income, attitude to health, tolerance of risk, etc. is driving our result we estimate a locally weighted polynomial regression over an outcome we hypothesize would be similarly affected by such a characteristic, but which is not correlated with access to sanitation. A plausible candidate is a non-hydric disease (a disease that is not transmitted by water), which is likely to be correlated with factors such as those mentioned but not with the access to sanitation. The outcome measure used was the prevalence of cough reported by the mother. Using a local polynomial regression, we estimate mean prevalence of cough at different village ratios of access, shown in Figure 9(a) and 9(b). We observe that there is no relationship between ratio of access to fixed-point defecation or improved sanitation and the prevalence of cough.

A second robustness check to test if there is some unobservable factor, such as income, explaining the ratio of access to sanitation is to substitute the rate of access to an alternative household infrastructure asset. We use the ratio of access to electricity, which is measured as the proportion of households who report having electricity. If we look at Figure 10, we can see that there is a flat relation when we estimate a local polynomial regression between the prevalence of diarrhea and the ratio of electricity.

When we estimate our parametric model but instead of controlling by the ratio of sanitation we incorporate the ratio of access to electricity at the village level (Figure 11(a) and 11(b)), we can see that there is a flat relation between the ratio of access to electricity and the probability of diarrhea prevalence.

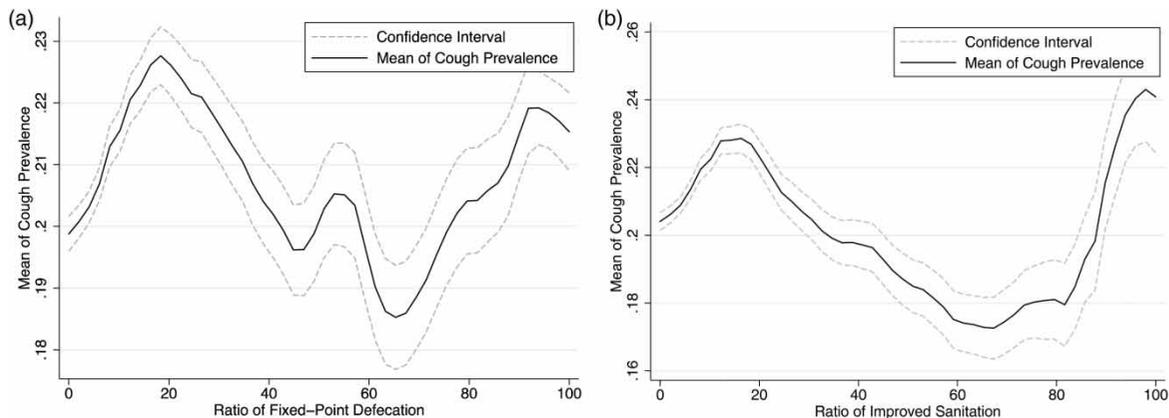


Figure 9 | Local polynomial regression cough and ratio of sanitation. (a) Fixed-point defecation. (b) Improved sanitation. *Note:* The curve presents the local polynomial regression between the prevalence of cough for each level of ratio of access to sanitation in the village. The pointed lines are the confidence interval at 95%.

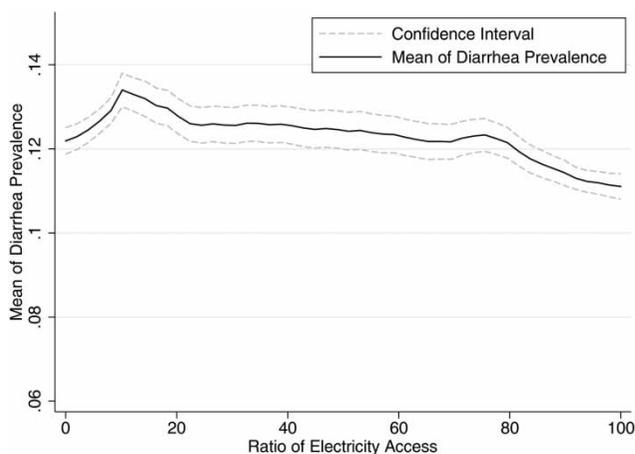


Figure 10 | Local polynomial regression diarrhea and ratio of electricity access. *Note:* The curve presents the local polynomial regression between the prevalence of diarrhea for each level of ratio of access to electricity in the village. The pointed lines are the confidence interval at 95%.

Counterfactual model

Finally, we carry out a series of non-parametric estimations to replicate and test the robustness of the results in Figure 6(a) and 6(b). We estimate a non-parametric matching estimator comparing children living in households with/without access to sanitation in villages with no coverage to children living in households with/without access in villages at different rates of access to sanitation. We use as a baseline benchmark $D_{ijk}(S_{jk}^v = 0, R_k^v = 0)$ for each node of the distribution of $D_{ijk}(S_{jk}^v = s, R_k^v = r)$ where $s \in \{0, 1\}$ and $s \in [0.1, 0.9]$. For each node, we

estimate an ATE shown in Equation (12) through matching estimators:

$$\begin{aligned} \frac{\partial D_{ijk}}{\partial (S_{jk}^v, R_k^v)} &= E [D_{ijk} | S_{jk}^v = s, R_k^v = (r, r + 10)] - E [D_{ijk} | S_{jk}^v = 0, R_k^v = (r^*, r^* + 10)] \end{aligned} \tag{12}$$

We estimate each of these nodes ($S_{jk}^v = s, R_k^v = (r, r + 10)$) using a nearest neighborhood matching estimator with two matches as controls for each observation in the treatment group. The results of these estimations are presented in Figure 12(a) and 12(b) for fixed-point defecation and improved sanitation, respectively. The graph demonstrates results similar to those estimated with our parametric model (Figure 6(a) and 6(b)) with an average direct benefit of 1 pp for access to fixed-point defecation and 1.5 pp for access to improved sanitation. When we look at the combined effect of going from a household without sanitation in a village without coverage to a household with fixed-point defecation in a village with complete coverage we find a benefit equivalent to a reduction in diarrhea prevalence of 3 pp. This is analogous to the difference in Figure 6(a) where the predicted probability of diarrhea for $E[D_{ijk} | S_{jk}^0 = 0, R_k^v = 0]$ is 12.5% and $E[D_{ijk} | S_{jk}^1 = 1, R_k^v = (80, 90)]$ is 1.0 pp, a reduction of 2.5 pp.

Similarly, the results for improved sanitation show a direct benefit that ranges from [0.5,2], which on average is consistent with the 1.4 pp reduction in diarrhea prevalence

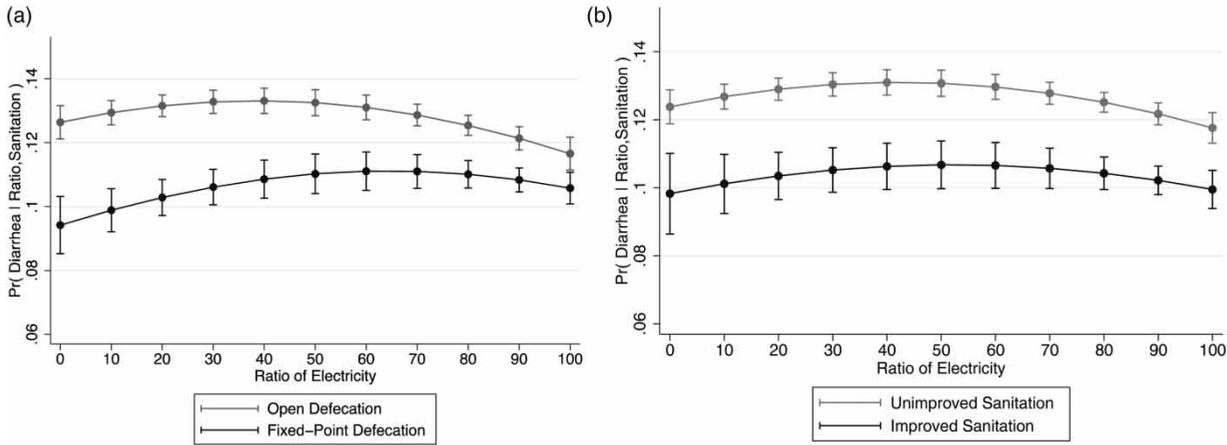


Figure 11 | Parametric model controlling with electricity ratio. (a) Fixed-point defecation. (b) Improved sanitation. Note: Calculated as $Pr(D_{ijk} | R_k^v, S_{jk}^v) |_{R_k^v=J}$ for each value of electricity access ratio, $J \in [0, 100]$. The bars at each point of estimation correspond to the confidence interval using robust clustered errors at the village level.

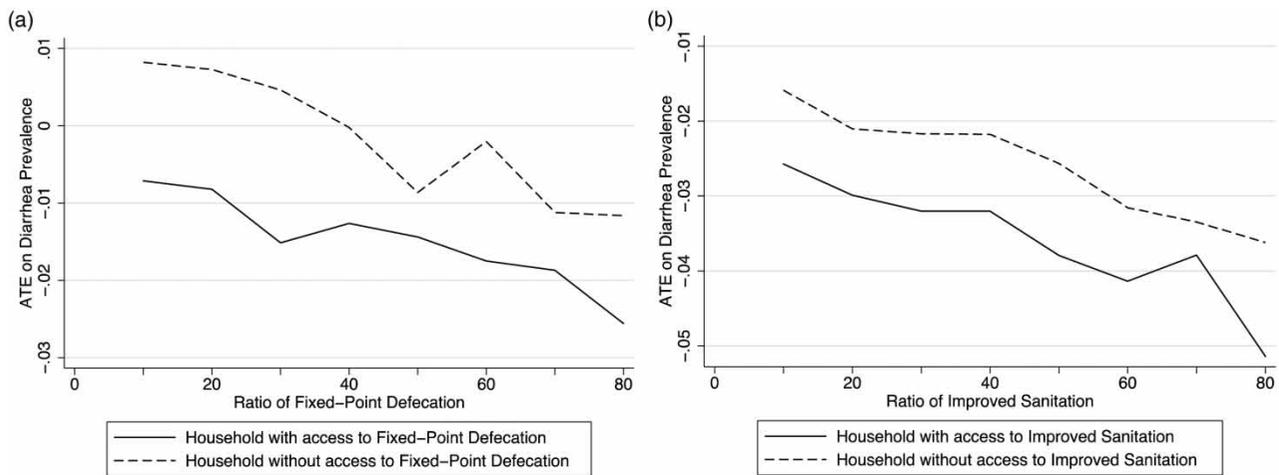


Figure 12 | Counterfactual model. (a) Fixed-point defecation. (b) Improved sanitation. Note: The parameters presented in the graph are calculated for each node (r, s) as $E[D_{ijk} | S_{jk}^v = s, R_k^v = (r, r + 10)] - E[D_{ijk} | S_{jk}^v = 0, R_k^v = (r, r + 10)]$. Where $r \in [10, 90]$, $s \in \{0, 1\}$. These parameters were estimated using a nearest neighborhood propensity score matching estimator with 30 neighbors. The variables used to estimate the propensity score for the matching are sex of the household, if the household head is from a low caste, the age of the child, and the sex of the child.

that we find in our parametric model. When we look at the difference between $E[D_{ijk} | S_{jk}^1 = 0, R_k^v = 0] = 12.5$ and $E[D_{ijk} | S_{jk}^1 = s, R_k^v = (80, 90)] = 5.1$ it is consistent with the parameter estimated with our counterfactual model for the node $(80, 1)$ which is a reduction in diarrhea prevalence of 5.1 pp.

CONCLUSIONS

This paper investigates the sources of health benefits to improved sanitation infrastructure access and elimination

of open defecation. We hypothesize that there are two sources of benefit, a direct benefit on intra-household child health when moving from open defecation to accessing sanitation, and an externality benefit driven by the ratio of access to sanitation at the village level. We estimate our parametric model of the treatment effect of access to sanitation on child health outcomes to test for both the marginal effect on the probability of diarrhea of sanitation access (direct benefit), as well as the change and the rate of change in diarrhea resulting from externalities to sanitation access. Consistent with other studies on this topic, we find evidence of positive direct benefits for both access

to fixed-point defecation and improved sanitation. On average, if a child moves from a household practicing open defecation or using an unimproved sanitation facility to a household with improved sanitation, there is a direct benefit of 10% (1.26 pp) reduction in diarrhea prevalence. However, we find that the external benefit from community coverage is more than three times the individual benefit, that is, an additional 37% (4.5 pp) reduction in diarrhea prevalence. We carry out a series of systematic robustness checks to test whether characteristics besides access to sanitation could be driving these results and find that our results hold under each of these alternative econometric specifications.

The magnitude of direct benefits is smaller than shown in previous studies, but this is consistent with our hypothesis that positive externalities to improved sanitation access drive much of the benefit of sanitation. More importantly, we find that the benefits from externalities follow a concave function in that benefits begin to accrue as higher level of community coverage is met, and these benefits accrue at a faster rate as coverage increases. This suggests that the goal of ODF communities as proposed by the post-2015 SDG working group on sanitation goals is socially optimal. For the full benefit of sanitation infrastructure to be realized efforts should focus on achieving community-wide coverage of improved sanitation and elimination of open defecation.

These results make two important contributions to the economics literature on health outcomes in early childhood – as an indicator of long-term human capital attainment – and the public policies that address inequalities in these outcomes. The first is that the results of this study support the idea of household sanitation as a public good with public benefits, rather than a private good with only private benefits. The existence of externalities to sanitation suggests that households will underinvest in sanitation and that private markets alone will fail to produce optimal outcomes. The second contribution is to demonstrate a potential role for public sector intervention, either by easing budget constraints or increasing access to information about the benefits of sanitation. Indeed, 77% of our sample of rural households lacks access to improved sanitation, suggesting that the potential health benefits from increasing access to sanitation are huge.

We can see from our sample that a great proportion of the rural population lives in villages with low sanitation coverage. However, our results show that most of the benefits from sanitation, especially their externalities, appear only at upper levels of coverage, giving us a feeling that the best is yet to come.

ACKNOWLEDGEMENTS

We are thankful to Olga Namen, Eduardo Perez, Payal Jayant, Dean Spears, and anonymous referees for their useful comments. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the view of the World Bank, its Executive Directors, or the countries they represent. The findings, interpretations, and any remaining errors in this paper are entirely those of the authors.

REFERENCES

- Andres, L., Iimi, A., Orfei, A. & Samad, H. 2013 *Impact Evaluation for Infrastructure: General Guidance and Existing Evidence*. The World Bank, Washington, DC.
- Bertrand, M., Duflo, E. & Mullainathan, S. 2004 *How much should we trust differences-in-differences estimates?* *The Quarterly Journal of Economics* **119** (1), 249–275.
- Bose, R. 2009 *The Impact of Water Supply and Sanitation Interventions on Child Health: Evidence from DHS Surveys*. International Initiative for Impact Evaluation, New Delhi, India.
- Cameron, L. & Shah, M. 2013 *Impact Evaluation of Rural Sanitation in East Java, Indonesia*. The World Bank, Water and Sanitation Program, Washington, DC.
- Checkley, W., Buckley, G., Gilman, R. H., Assis, A. M. O., Guerrant, R. L., Morris, S. S., Mølbak, K., Valentiner-Branth, P., Lanata, C. F., Black, R. E., Childhood Malnutrition and Infection Network 2008 *Multi-country analysis of the effects of diarrhoea on childhood stunting*. *International Journal of Epidemiology* **37** (4), 816–830.
- Chen, L. C., Huq, E. & Huffman, S. L. 1981 *A prospective study of the risk of diarrheal diseases according to the nutritional status of children*. *American Journal of Epidemiology* **114** (2), 284–292.
- Clasen, T., Boisson, S., Routray, P., Torondel, B., Bell, M., Cumming, O., Ensink, J., Freeman, M., Jenkins, M., Odagiri, M., Ray, S., Sinha, A., Suar, M. & Schmidt, W.-P. 2014 *Effectiveness of a rural sanitation programme on diarrhoea*,

- soil-transmitted helminth infection, and child malnutrition in Odisha, India: a cluster-randomized trial. *The Lancet Global Health* **2** (11), e645–e653.
- Das, J., Hammer, J. & Sánchez-Paramo, C. 2012 The impact of recall periods on reported morbidity and health seeking behavior. *Journal of Development Economics* **98** (1), 76–88.
- Duflo, E., Greenstone, M., Guiteras, R. & Clasen, T. 2015 *Toilets Can Work: Short and Medium Run Health Impacts of Addressing Complementaries and Externalities in Water and Sanitation*. Tech. Rep. 9, National Bureau of Economic Research. <http://www.nber.org/papers/w21521>.
- Eisenberg, J. N. S., Desai, M. A., Levy, K., Bates, S. J., Liang, S., Naumoff, K. & Scott, J. C. 2007 Environmental determinants of infectious disease: a framework for tracking causal links and guiding public health research. *Environmental Health Perspectives* **115** (8), 1216–1223.
- Eisenberg, J. N. S., Trostle, J., Sorensen, R. J. D. & Shields, K. F. 2012 Toward a systems approach to enteric pathogen transmission: from individual independence to community interdependence. *Annual Review of Public Health* **33**, 239–257.
- Fan, V. Y.-M. & Mahal, A. 2011 What prevents child diarrhoea? The impacts of water supply, toilets, and hand-washing in rural India. *Journal of Development Effectiveness* **3** (3), 340–370.
- Fewtrell, L., Kaufmann, R. B., Kay, D., Enanoria, W., Haller, L. & Colford Jr, J. M. 2005 Water, sanitation, and hygiene interventions to reduce diarrhoea in less developed countries: a systematic review and meta-analysis. *The Lancet Infectious Diseases* **5** (1), 42–52.
- Gersovitz, M. & Hammer, J. S. 2004 The economical control of infectious diseases. *The Economic Journal* **114** (492), 1–27.
- Geruso, M. & Spears, D. 2015 *Neighborhood Sanitation and Infant Mortality*. Technical report, National Bureau of Economic Research. <http://www.nber.org/papers/w21184>.
- Gunther, I. & Fink, G. 2010 *Water, Sanitation and Children's Health: Evidence from 172 (DHS) Surveys*. Policy Research Working Paper Series, No. 5275. World Bank.
- Gupta, A., Spears, D., Coffey, D., Khurana, N., Srivastav, N., Hathi, P. & Vyas, S. 2014 Revealed preference for open defecation. *Economic and Political Weekly* **49** (38), 43–55.
- Headey, D. D. 2013 Developmental drivers of nutritional change: a cross-country analysis. *World Development* **42**, 76–88.
- Humphrey, J. H. 2009 Child undernutrition, tropical enteropathy, toilets, and handwashing. *Lancet* **374** (9694), 1032–1035.
- Hunter, P. R. & Prüss-Ustün, A. 2016 Have we substantially underestimated the impact of improved sanitation coverage on child health? A generalized additive model panel analysis of global data on child mortality and malnutrition. *PLoS One* **11** (10), e0164571.
- International Institute for Population Sciences 2010 *District Level Household and Facility Survey (DLHS-3), 2007–08*. Technical report, International Institute for Population Sciences, Mumbai, India.
- JMP 2012 *Progress on Drinking Water and Sanitation: 2012 Update*. Technical report, WHO/UNICEF Joint Monitoring Programme for Water Supply and Sanitation, New York, USA.
- JMP 2015 *Progress on Drinking Water and Sanitation: 2013 Update*. Technical report, WHO/UNICEF Joint Monitoring Programme for Water Supply and Sanitation, New York, USA.
- Kumar, S. & Vollmer, S. 2012 Does access to improved sanitation reduce childhood diarrhea in rural India? *Health Economics* **22**, 410–427.
- Leuven, E. & Sianesi, B. 2012 PSMATCH2: Stata Module to Perform Full Mahalanobis and Propensity Score Matching, Common Support Graphing, and Covariate Imbalance Testing. Statistical Software Components. Available from <http://econpapers.repec.org/software/bocbocode/s432001.htm>.
- Liu, L., Johnson, H. L., Cousens, S., Perin, J., Scott, S., Lawn, J. E., Rudan, I., Campbell, H., Cibulskis, R., Li, M., Mathers, C. & Black, R. E. 2012 Global, regional, and national causes of child mortality: an updated systematic analysis for 2010 with time trends since 2000. *The Lancet* **379** (9832), 2151–2161.
- Moulton, B. R. 1990 An illustration of a pitfall in estimating the effects of aggregate variables on micro units. *The Review of Economics and Statistics* **72**, 334–338.
- Nadaraya, E. A. 1964 On estimating regression. *Theory of Probability & Its Applications* **9** (1), 141–142.
- Patil, S. R., Arnold, B. F., Salvatore, A. L., Briceño, B., Ganguly, S., Colford Jr, J. M. & 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 Med* **11** (8), e1001709.
- Prüss, A., Kay, D., Fewtrell, L. & Bartram, J. 2002 Estimating the burden of disease from water, sanitation, and hygiene at a global level. *Environmental Health Perspectives* **110** (5), 537–542.
- Spears, D. 2012 *Effects of Rural Sanitation on Infant Mortality and Human Capital: Evidence from India's Total Sanitation Campaign*. Technical report, RICE working paper.
- Spears, D. 2013 *How much International Variation in Child Height can Sanitation Explain?* Policy Research Working Paper. no. WSP 6351.
- Spears, D., Ghosh, A. & Cumming, O. 2013 Open defecation and childhood stunting in India: an ecological analysis of new data from 112 districts. *PLoS One* **8** (9), e73784.
- UNICEF 2012 *Committing to Child Survival: A Promise Renewed*. Progress Report 2012.
- Victora, C. G., Adair, L., Fall, C., Hallal, P. C., Martorell, R., Richter, L. & Sachdev, H. S. 2008 Maternal and child undernutrition: consequences for adult health and human capital. *Lancet* **371** (9609), 340–357.
- Watson, G. S. 1964 Smooth regression analysis. *The Indian Journal of Statistics* **26** (4), 359–372.