GEORGETOWN UNIVERSITY
McCourt School of Public Policy

Measuring the Effect of iDE's Sanitation Marketing Intervention on Latrine Coverage and Health in Cambodia

Mutiara Alam Addini
Cassandra Ake
Anna Konstantinova
Akaravuit Pancharoen

April 27th, 2018

Table of Contents

List of Acronyms

ANCOVA	Analysis of Covariance
CMRD	Cambodian Ministry of Rural Development
CSES	Cambodia Socio-Economic Survey
DALY	Disability Adjusted Life Year
DHS	Demographic Health Survey
IDE	International Development Enterprises
JMP	Joint Monitoring Programme
QGIS	Quantum Geographic Information System
SD	Standard Deviation
SDG	Sustainable Development Goals
SMSU	Sanitation Marketing Scale-Up Program
UN	United Nations
UNICEF	United Nations International Children's Fund
USAID	United States Agency for International Development
WASH	Water Sanitation and Hygiene
WHO	World Health Organization
WFP	World Food Programme

List of Tables

List of Figures

I. Executive Summary

I.A. Background

Access to improved water and sanitation is a major development challenge for Cambodia. The majority of Cambodia's rural poor still have relatively low access to improved water and sanitation, low water quality, and 8.6 million continue to practice open defecation. Healthcare experts believe these poor sanitation and hygiene practices lead to compromising health conditions, such as diarrhea and parasitic infections, which result in high mortality rates for children under-5 and have long-run impacts on their quality of life.

To improve the sanitation status quo in rural Cambodia, International Development Enterprises (iDE) developed the Sanitation Marketing Scale-Up (SMSU) program. This program has three broad aims: (1) to encourage latrine producers to produce improved (pour-flush) latrines; (2) to educate communities on the benefits of improved latrines, and; (3) to mediate sales between households and latrine producers. The SMSU 1.0 program ran from August 2011 to October 2014 facilitating sales of over 141,000 latrines in rural areas, although 37% of these were uninstalled nine months after delivery.

I.B. Motivation and Research Question

International Development Enterprises (iDE) invited the Capstone Team to assist in building an econometric model that captures their latrine market intervention's effect on increasing latrine coverage and improving health effects such as diarrhea for children under-5. Specifically, they sought to answer three questions: (1) did iDE's intervention lead to increases in latrine coverage; (2) did iDE's intervention lead to improvements in health, and; (3) what is the relationship between latrine coverage and health? The final question was further decomposed into three parts: whether having a latrine is associated with healthier children in the household; whether associated health effects differ depending on the presence of the other sanitation-related disease transmission pathways; and whether the associated changes to children's health are affected by village latrine coverage rates.

A literature review showed that the most rigorous experimental studies of latrine interventions are inconclusive about the health effects of increasing latrine coverage and have been unable to replicate the overwhelmingly positive findings in older studies.[1] In addition, diarrhea was most often the health effect of interest and other known health effects such as malnutrition, stunting, wasting, helminth infection, anemia, and roundworm infestation have received less attention. The literature also suggested that the health benefits derived from latrines can be affected by community practices regarding shared latrines, community-level latrine coverage, the presence of livestock, the type of

[1] Experimental studies include Clasen et al. (2014), Patil et al. (2015), Pickering et al. (2015), Briceño et al. (2017), and Cameron and Shah (2017). Prior epidemiological studies are summarized in systematic reviews by Esrey et al. (1985), Esrey et al. (1991), Waddington et al. (2009), Fewtrell et al. (2005), and Wolf et al. (2005).

dwelling one lives in, access to clean drinking water and handwashing facilities, household income, and parental educational attainment.[2]

I.C. Methodology

The data available for this analysis included primary data collected by iDE during its annual village-level Latrine Count Survey and secondary data from the Cambodia Socio-Economic Survey (CSES) and the USAID-funded Demographic Health Survey (DHS). Given the purpose of each of these surveys and their level of correlation with iDE's data, we conclude CSES is the most reliable in terms of its latrine coverage and socioeconomic measures, while DHS is most reliable for health measures such as diarrhea prevalence.

Our report addresses the three research questions by employing a range of evaluation methods. We explore question 1 by using a difference-in-difference model to measure iDE's causal impacts on changes in latrine coverage at the commune- and district-levels using both CSES and DHS data. The difference-in-difference estimation is also used to answer question 2 and measure the effect of iDE's program on health outcomes at the district level using DHS data. For question 3, we first use a fixed effects model to try to uncover a causal relationship between latrines and health using DHS data. We then further characterize the relationship between latrines and health using several multivariate regression models employing iDE, CSES, and DHS data.

I.D. Results

Our analyses of iDE's intervention result in the following conclusions:

1. Using commune-level CSES data, we find that iDE's intervention has led to increases in rural latrine coverage compared to other regions of the country. Difference-in-difference analysis shows that latrine coverage increases by 17.8 percentage points, or a 94% increase, in iDE target areas after the intervention. However, these effects are not replicated when using district-level DHS data.[3]

2. Using district-level DHS data, difference-in-difference analysis shows that after IDE's intervention, rural diarrhea prevalence in iDE target areas decreased by 5.8 percentage points. Due to measurement inconsistences, the CSES data could not be used to verify these results. In addition, we do not see any effects on other health outcomes such as malnutrition, stunting, and wasting, but the short duration (two years) between implementation and evaluation may be insufficient to observe changes on those outcomes. Preliminary analysis of whether the

[2] Studies include Clasen et al. (2012), Cairncross (1990), Yishay et al. (2017), Mosites et al. (2015), Heijnin et al. (2014), Clasen et al. (2006), Exum et al. (2016), Cattaneo et al. (2009), and Ejemot et al. (2008).

[3] There are approximately 100 communes in each district. As a result, the sample size in the commune-level CSES analysis was 1,124 whereas the sample size for the district-level DHS analysis was only 323. The smaller sample size in the DHS analysis reduces statistical power and makes it difficult to detect treatment effects.

observed health effects vary in the presence of other sanitation pathways and socioeconomic factors suggest exposure to iDE's intervention leads to greater improvements in health when combined with higher rates of using a clean drinking water source and treating one's drinking water. However, low statistical power prevents us from drawing strong conclusions on these effects and from determining whether other factors like handwashing facilities, education, income, livestock ownership, or concrete floors had similar effects.

3. The fixed effects model used with the DHS dataset failed to find any statistically significant evidence of a causal relationship between latrine coverage and changes to health. When using IDE's own data and non-causal analysis, we found a strong negative association between latrine coverage and their diarrhea indicator. Controlling for socioeconomic and other sanitation factors, a 10 percentage point increase in village latrine coverage is associated with a 5.7 percentage point reduction in the fraction of households in the village where at least one person had diarrhea in the last month. Although analysis at the household level using the DHS and CSES datasets did not show a statistically significant association between latrine ownership and diarrhea prevalence, there was a strong relationship between improved physical growth measures and a household having both a latrine and at least one other disease transmission barrier such as handwashing facilities, clean drinking water sources, drinking water treatment, or reduced livestock levels. The effects of shared latrines, village-level latrine coverage rates, and whether the household had concrete floors are inconclusive.

I.E. Limitations and Policy Implications

Our findings on the impact of iDE's sanitation marketing intervention are generally positive but the study's internal validity is challenged by three limitations. First, each of the surveys were designed to achieve different goals, which affected the construction and accuracy of the variables used in our analysis. Second, these surveys were designed to estimate national and provincial-level statistics. Using the provided provincial weights to construct the commune and district averages used in our analysis may have introduced bias, though it is difficult to determine the direction of the bias a priori. Finally, there is little overlap between the outcomes measured in the primary and secondary data sources and data was not available from all three sources in the same year to capture behavior prior to the intervention in 2012.

In terms of policy implications, our study has shown that market-based interventions are effective in terms of increasing latrine coverage. This finding is robust to the inclusion of community-level characteristics that may affect knowledge of proper sanitation practices and the ability to purchase a latrine. Thus, the primary goal of iDE's intervention has been achieved: more people have access to improved sanitation.

These market-based interventions may have resulted in improvements in health, but given the multiple steps between purchase and changes to well-being, the relationship is far from straightforward. First, a latrine must be used consistently for any health effects to be observed. Our analysis also suggests that a latrine may be insufficient to change non-diarrheal health outcomes without additional protective sanitation and hygiene practices. Despite our initial findings, further data is needed to study the effect of latrines in the context of other sanitation practices. Some of these intervening pathways may matter more or less depending on whether one is examining effects at the household-, village-, commune-, or district-level; for example, reaching a particular threshold of latrine coverage might only relevant be at the village-level. Practitioners should think carefully about these relationships when designing their intervention's theory of change and monitoring and evaluation frameworks.

Moreover, WASH practitioners designing surveys to capture household- and community- level characteristics as well as health impacts should think carefully about their methodology. For example, we have seen that conducting censuses are more accurate when doing econometric analysis at lower geographic levels compared to relying on sampling means. In addition, a specific description of what constitutes diarrheal incidence is necessary for household respondents to answer with accuracy. Recall period is also an important factor in receiving reliable responses, as we saw from the differing variability in responses to whether someone had diarrhea when different recall periods were used. Our data suggests two weeks is likely the most accurate.

In addition, further research is needed to firmly establish the causal relationship between increased latrine coverage and improved health impacts. Several of these health impacts may only emerge after many years have passed and may be difficult to identify depending on the geographical level at which the analysis takes place. In addition, more research is needed to evaluate whether latrine interventions should be coupled with other water and sanitation interventions, and whether behavior change campaigns are necessary. Finally, some health impacts, such as diarrhea, may vary according to season, but data often lacks the level of temporal identification needed to measure seasonal variability.

By concentrating on improved latrine coverage, Cambodia can move toward achieving the UN's Sustainable Development Goals (SDG) 3 and 6, thereby ensuring the health of all of its children. iDE's intervention has been shown to be effective in achieving greater household access to latrines and we would encourage them and the Government of Cambodia to work together to ensure the proper level of support, regulation, and financing in order to further expand latrine coverage in the country.

II. Introduction

International Development Enterprises (iDE) invited the Capstone Team to assist in building an econometric model that captures their latrine market intervention's effect on increasing latrine coverage and improving health outcomes for rural Cambodian children under-5. Specifically, they sought to answer three questions: 1) did iDE's intervention lead to increases in latrine coverage; 2) did iDE's intervention lead to improvements in health, and; 3) what is the relationship between latrine coverage and health?

This report details the quantitative study designed to answer these questions, conducted using iDE's own data and two secondary data sources. To give context to the importance of these research questions to Cambodia's development, Section III details the state of sanitation in Cambodia and how iDE's Sanitation Marketing Scale-Up program is attempting to address these issues. The section also provides a research motivation and a brief discussion of the literature, which contributed to a theory of change to relate increases in latrine coverage to improved health outcomes. A more exhaustive discussion of the literature is presented separately.

Section IV provides a description of the primary and secondary data sources used and the construction of the outcome variables of interest. Section V details the empirical strategy employed to answer each of the above research questions and includes the estimating equations used. The results of the empirical strategy are in Section VI, with a separate subsection devoted to each of the research questions listed above. Finally, Section VII concludes while providing a brief discussion of the data limitations and the policy ramifications of our findings.

III. Background and Context

III.A. The State of Sanitation in Cambodia

Ensuring access to water and sanitation for all is one of the main themes of the United Nations' development goals, included under both the Millennium Development Goals and the Sustainable Development Goals. Unfortunately, Cambodia currently has the lowest coverage of improved water and sanitation facilities in Southeast Asia (Chase, 2015). According to the World Health Organization in 2010, 8.6 million Cambodians continue to practice open defecation, a sanitation gap that would require approximately 1.2 billion USD to close (Hutton, 2012).

Access to improved sanitation in the country is highly dependent on wealth (Smets, 2015). Only 12% of the poorest income quintile of the rural population have access to improved sanitation, whereas 59% of the richest rural quintile have improved latrines (Smets, 2015). 66% of Cambodian households had access to improved water sources in 2012, but the quality of water continues to be a concern (Smets, 2015). While knowledge of the importance of handwashing for the prevention of diarrhea is high at 72%

(CMRD, 2010), 41% of rural Cambodians had limited or no access to handwashing facilities in 2015 (UNICEF, 2017).

This continued sanitation gap, particularly in rural areas, has a profound impact on health. There are many diseases associated with poor sanitation and hygiene practices, among them diarrhea, dysentery, cholera, parasitic diseases, and other maladies (World Bank, 2008b). These diseases not only generate costs in terms of healthcare and productivity, but they also lead to high rates of death in children under-5. The World Bank estimates that diarrhea accounts for 19% of deaths in children under-5 in Cambodia, whereas malnutrition resulting from poor sanitation accounts for 18% of deaths (World Bank, 2008a). Although not always leading to death, diarrhea, malnutrition, and these other conditions have long-run impacts on quality of life. Adding to the dire state of child health in the country, nearly 40% of Cambodian children are stunted, over 28% are underweight, and nearly 11% are severely malnourished (WFP, 2017).

III.B. Background of iDE's Sanitation Marketing Scale-Up Program

To improve the sanitation status quo in Cambodian rural areas, International Development Enterprises (iDE) developed the Sanitation Marketing Scale-Up (SMSU) program. iDE believes in the power of the market to solve development issues and works in 11 countries worldwide, primarily on WASH and agriculture projects. The SMSU program in Cambodia encourages latrine producers to produce improved (pour-flush) latrines, educates communities on the benefits of improved latrines, and mediates sales between households and latrine producers. This approach, known as Sanitation Marketing, believes that by treating poor households as consumers rather than as welfare beneficiaries, private producers will develop latrine models that the poor will be more interested in purchasing and using (Dumpert and Perez, 2015). In this case, iDE assisted private producers by funding prototype development using human-centered design techniques to understand consumers' preferences and the issues they face.

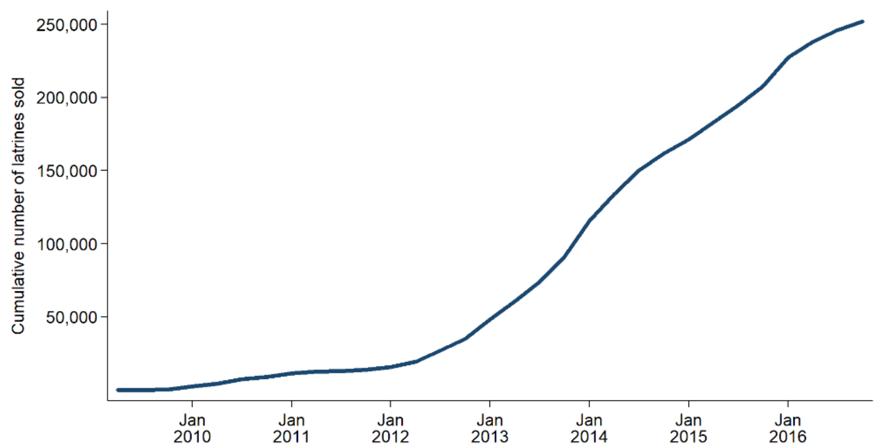
To educate communities about the benefits of latrines, sanitation sales agents conduct group meetings and door-to-door presentations that help rural households weigh the costs of buying a latrine versus continuing to defecate in open fields. They then enter the sales information into a smartphone app, and orders are distributed to available local business owners to try to eliminate long customer wait times. These sanitation sales agents are paid a fraction of the profits made by local latrine business owners for each sale. In addition, in Kandal and Svay Rieng provinces, 20,000 individuals were subject to a more intense behavior change campaign that incorporated local beliefs about propriety, night-time safety when using a latrine, and latrine affordability.[4] A pilot program in Kandal and Prey Veng provinces also tested a sanitation financing project, since micro-finance institutes

[4] https://s3.amazonaws.com/www.ideglobal.org/files/public/iDE-Expanding_Reach_Sanitation_Market_BCC_Cambodia.pdf?mtime=20160621232040

are often unwilling to finance lower priced, non-income-generating assets. This pilot program found that 37.8% of households used financing to purchase their latrine.[5]

To date, the intervention has resulted in over 250,000 latrines sold in rural areas; 141,000 of these were sold by the end of 2014, the year we treat as post-intervention in our analysis. This increase in latrine sales is represented in Figure 1 below. According to the SMSU 1.0 final report, the project reached a 45% latrine coverage rate in the seven provinces where iDE works, representing a significant increase of 16 percentage points in coverage over just 2.5 years. The success of SMSU 1.0, which ran from August 2011 to October 2014, is the subject of this study. However, the project scaled up operations more fully by 2012, so 2012 is considered the baseline for purposes of our analysis.

Figure 1: Cumulative iDE latrine sales over time.



III.C. Research Motivation

Having achieved rapid increases in latrine coverage in rural areas, iDE is grappling with the question of whether their program has had an impact on health through changes in diarrhea prevalence, among other possible health outcomes. While the organization believes it has increased latrine coverage in its target areas, coverage does not necessarily translate to usage, a linkage that is vital if health is to improve. iDE hypothesizes in its programmatic reports that a usage gap could result from a number of factors, such as the additional cost of installation and the consumers' preference for a shelter to allow for privacy. This shelter is not included in the purchase of a latrine. Through monitoring, iDE also noticed that 37% of latrines remained uninstalled nine months after delivery. The program was subsequently modified (in March 2015) so that latrine prices included installation services to address this gap. Other possible barriers to usage include the difficulty for very young children to use the facilities, latrine unavailability when working away from home, or the cost of emptying the latrine when it fills.

[5] https://s3.amazonaws.com/www.ideglobal.org/files/public/iDE-FR_WASH_SMSU_Cambodia.pdf?mtime=20161013212546

As discussed in detail in the literature review, experimental studies are inconclusive about the health effects of increasing latrine coverage and have been unable to replicate the overwhelmingly positive findings in older literature employing less rigorous research methods. While copious literature exists about the relationship between diarrhea and latrines, other health effects such as malnutrition, stunting, wasting, anemia, helminth infection, and roundworm infestation have received less attention.

The older literature referenced above generally consisted of large epidemiological studies using national surveys of health outcomes. These epidemiological studies showed some association between latrine coverage and decreased rates of diarrhea despite considerable heterogeneity and methodological shortcomings (Kumar and Vollmer, 2012; Vásquez and Aksan, 2015; Clasen et al., 2010; Schmidt, 2014; Waddington et al., 2009). Mirroring the results found in epidemiological research, systematic reviews of the most rigorous of the quasi-experimental research on the effects of latrine interventions show average reductions in diarrheal morbidity of 36 percent (Esrey et al., 1985; Esrey et al., 1991; Waddington et al., 2009; Fewtrell et al., 2005; Wolf et al., 2005). However, recent rigorous studies using randomized evaluations have found few to no causal linkages between latrine interventions and decreased diarrhea prevalence (Clasen et al., 2014; Patil et al., 2015; Pickering et al., 2015; Briceño et al., 2017; Cameron and Shah, 2017).

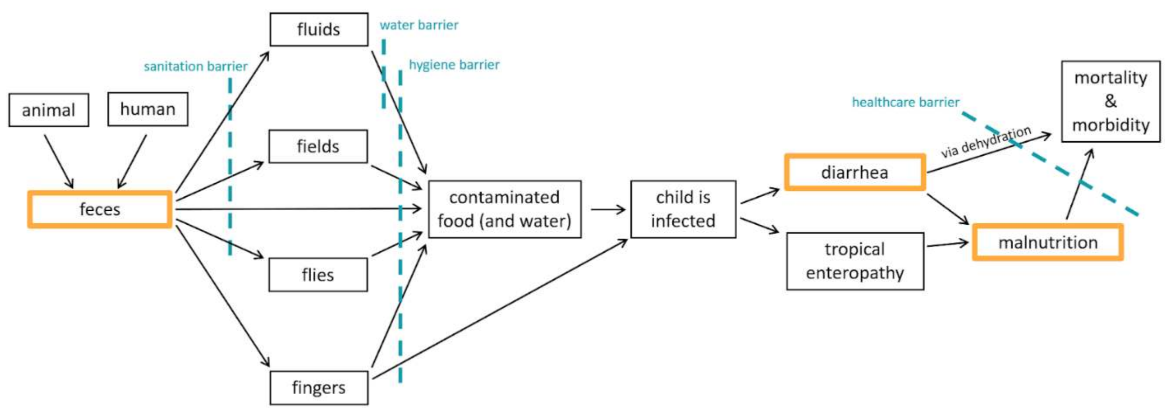
In terms of other health effects, Rah et al. (2015)'s epidemiological study found that households in India with access to a latrine have a 14% lower prevalence of stunting among their children. Particularly relevant to our setting, Kov et al. (2013) and Vyas et al. (2016) both find a relationship between improved sanitation and height-for-age z-scores (a measure of wasting) among Cambodian children. Moving on to the quasi-experimental evidence, Ziegelbauer et al. (2012) focused on four specific soil-transmitted helminth infections and found eleven studies which collectively suggest that having an improved latrine is associated with a 40 to 60 percentage point reduction in the odds of infection. However, the more rigorous randomized controlled trials found no results when examining health effects such as soil-transmitted helminth infections, weight-for-age z-scores, height-for-age z-scores, or anemia, with the exception of one study that found a 44% decrease in roundworm infestation following a latrine intervention (Cameron and Shah, 2017).

Facing the contradictions in the literature, iDE has tasked this Capstone Team with assembling data and developing an empirical strategy that can capture the health effects of iDE's latrine market intervention program in rural Cambodia. In order to prove that iDE's Sanitation Marketing Campaign has not only been successful in expanding latrine coverage but also in improving health impacts, the overall research question for the capstone project is: What was iDE's latrine marketing intervention's effect on increasing latrine coverage and improving health outcomes for rural Cambodian children under-5? By asking the research question, we are looking to identify a relationship between sanitation and better health outcomes, as well as how iDE's specific sanitation marketing approach can influence this relationship. Focusing this research question on health outcomes for children under-5 as the main population of interest is common in the literature, as they are the population most vulnerable to sanitation-related diseases (WHO/UNICEF Joint Monitoring Program, 2017).

III.D. Theory of Change

In order to draw a causal link between iDE's intervention and subsequent health impacts, it is necessary to consider the intervention within an explicit theory of change. To better understand how a latrine might prevent diarrhea and malnutrition, Figure 2 presents a model of the causal pathways linking poor sanitation (the absence of an improved latrine) to morbidity, meaning disease or illness, and mortality outcomes.

Figure 2: Detailed causal mechanisms connecting poor sanitation and under-5 mortality and morbidity. Pathways are adapted from the standard F-diagram of disease transmission and additional literature (Cheng et al., 2012; Clasen et al., 2012; Humphrey, 2009).



The left side of the diagram in Figure 2 depicts the standard epidemiological theory of disease transmission, often referred to as the F-diagram (Clasen et al., 2012). In this representation, human and animal feces can infect a person through the four Fs: fluids, fields, flies, and fingers (Clasen et al., 2012). Transmitting pathogens via fluids means the water used for dishes, cooking, and/or consumption contains bacteria or other organisms, whereas fields refer to the ground itself, which young children come into contact with more frequently than adults (Clasen et al., 2012). Flies act as a mechanical vector moving bacteria from one location to another and fingers refer to the cleanliness of hands, typically related to person-to-person contact (Clasen et al., 2012).

The right side of Figure 2 presents our own interpretation of the ways in which fecal infection leads to eventual morbidity or mortality, based on the available literature. The medical community broadly agrees that diarrhea leads to mortality through either (or both) dehydration and malnutrition (Cheng, 2012). The scarcity of medical records in most developing countries, where the most deaths due to diarrhea occur, prevents one from making a strong conclusion as to the preferential effect of dehydration versus malnutrition (Cairncross and Bartram, 2010). As a result, many sanitation programs opt for a mechanism-free way to measure impact by simply determining whether diarrhea prevalence has decreased or stayed the same.

Yet, research conducted by Humphrey et al. (2009) suggests that a focus on diarrhea prevalence might entirely miss the effect that contaminated food and water has on malnutrition through tropical enteropathy. Tropical enteropathy, though minimally studied, is a condition in which repeated exposure of the intestines to fecal matter causes a permanent reduction in the size of the intestinal ville (Waddington et al., 2009). Once that occurs, the body can absorb fewer nutrients through the intestine, leading to malnutrition, which results in poor physical growth outcomes like stunting, wasting, and malnourishment. Diarrhea may then exacerbate the low nutrient intake caused by tropical enteropathy, but current science is unable to distinguish between whether diarrhea or enteropathy is more responsible for malnutrition (Humphrey et al., 2009). As a result, a strict focus on diarrhea may strongly underestimate or completely miss the decrease in under-5 mortality caused by improved sanitation.

Even more than mortality, the effect of poor sanitation on morbidity, meaning disease or illness, must be considered in our study. Globally, lack of access to clean water, improved sanitation, and proper hygiene amount to 82,196,000 disability-adjusted life years (DALYs) per year, or ~6% of the total disease burden (Prüss et al., 2002). Given the many possible causal pathways between improper sanitation and mortality and morbidity, this study will therefore use a number of different health outcome indicators, including diarrheal prevalence and malnutrition indicators (weight-for-age, height-for-age, and weight-for-height z-scores, and stunting, wasting, and malnourishment prevalence). Given the short time frame covered by the study and the small sample size, changes in mortality would be difficult to capture and so are excluded.

III.E. Mediating Community Characteristics

Analyzing the F-diagram more closely, one sees several transmission pathways that would be unaffected by changes to sanitation. The installation of a latrine does not affect the water and hygiene barriers shown in the diagram, allowing infection to continue through contaminated water sources and dirty hands. As a result, the installation of an improved latrine may have a zero or highly variable health impact depending on the broader environment in which the child lives. Moreover, there are several factors that affect an individual household's ability to purchase a latrine or that can mediate the degree to which the sanitation barrier is effective in improving health outcomes. These include household income, parental educational attainment, household hygiene practices and drinking water sources, the presence of livestock, the type of dwelling one lives in, community practices regarding shared latrines, and community-level latrine coverage. These factors and the evidence available on their differential effects will be discussed next.

The first of these characteristics is the level of parental education and income, high levels of which we theorize will increase the likelihood of buying a latrine and engaging in additional good health and hygiene practices (Cairncross, 1990; Yishay et al., 2017). In addition, we hypothesize that improved sources of drinking water and treating one's drinking water will strengthen the water barrier shown in the theory of change model above, thereby weakening the "fluids" mode of transmission (Clasen et al., 2006). Moreover, regarding floor materials' role in disease transmission, children under the age of 5 can often be found playing on the floor, and dirt floors could hold a number of bacteria and parasites carried in on family members' shoes (Exum et al., 2016). A concrete floor would prevent some of this and weaken the "fingers" pathway that is activated when children play on the floor and then stick their fingers in their mouths (Cattaneo et al., 2009). However, if children wash their hands frequently, this could compensate for floor type and weaken the "fingers" transmission pathways, suggesting the presence of handwashing facilities may be important (Ejemot et al., 2008). Finally, even if a household made the proper choices for themselves, they will continue to be affected by their neighbors' open defecation practices, and by animals that either they or their neighbors own. While we hypothesize that increased exposure to animals will increase diarrhea incidence, it may improve malnutrition measures due to increased consumption of animal products (Mosites et al., 2015).

Individual household sanitation practices and socioeconomic characteristics are not the only factors that could affect disease transmission. This study must also consider the spillovers resulting from overall community sanitation access and open defecation levels. If the population is particularly dense and there is not universal latrine coverage, it is possible that open defecation could still infect a given child whose own family owns a latrine. Conversely, lowering open defecation rates by increasing community-wide latrine coverage could provide herd immunity[6] against infections for all, regardless of personal latrine ownership. In the absence of full private latrine coverage, public or shared latrines could also cut down on open defecation rates and protect the community against some infections, but they may also lead to increased infection transmission (Heijnen et al., 2014). In fact, the WHO-UNICEF Joint Monitoring Programme (JMP) continues to define shared latrines as an unimproved sanitation source (WHO/UNICEF JMP, 2017). This study, therefore, will explore heterogenous effects at different village latrine coverage threshold levels while also comparing households using a private latrine with households using shared or public facilities.

Taking this set literature into account, the econometric models estimating the effect of iDE's intervention on health and the relationship between latrines and health will include variables capturing various aspects of household community practices regarding shared latrines, community-level latrine coverage, the presence of livestock, the type of dwelling one lives in, household income, household hygiene and drinking water practice, and parental educational attainment.

IV. Data

Our analysis draws together primary data collected by iDE and secondary data available from the Cambodia Socio-Economic Survey (CSES) and the Demographic Health Survey (DHS). A brief description of each data source and the approach used to construct the datasets used in the analysis follows.

[6] Herd immunity is defined as the resistance to the spread of contagious disease within a population that results if a sufficiently high proportion of individuals are immune to the disease, especially through vaccination. While diarrhea is not a contagious disease and there is no vaccine available, the overall pathogen load in a community will theoretically decrease as fewer individuals practice open defecation.

IV.A. Description of Data Sources

iDE has conducted the latrine count survey in its seven intervention provinces annually from 2012 to 2014 during the first iteration of its SMSU program. Another round of latrine count data was collected in 2016 during the program's second iteration. The primary sampling unit is the village. Villages were chosen using a two-stage stratified sampling strategy focused only on rural areas. Within each village, a census of all households was completed and village-aggregate results were reported. This survey methodology provides the most accurate measure of latrine coverage at the village level compared to the other data sources used in our analysis, which rely on sampling to determine village latrine coverage rates. In 2016, a question was added to the survey wherein households were asked if anyone living there had experienced diarrhea in the last month. This question did not separately identify children under-5 and did not specify the number of household members who experienced diarrhea. Although this outcome will be referred to as diarrhea prevalence in the results described in Section VI.C., it is not an actual measure of prevalence which requires knowing how many individuals experienced diarrhea compared to the entire population.

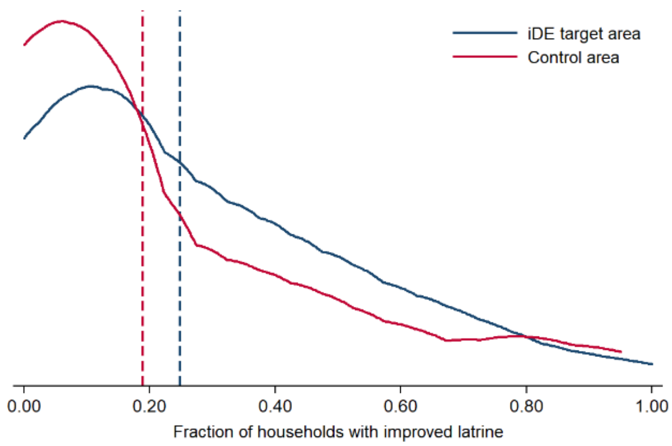
The Cambodia Socio-Economic Survey (CSES) is conducted by the Cambodian National Institute of Statistics and data was obtained for 2004, 2009, and 2014. The survey was conducted using a three-stage stratified design. In the first stage, the country was divided into rural-urban strata within each province and villages were randomly chosen from each stratum. In the second stage, each village was divided into enumeration areas and a single enumeration area was randomly chosen from each village. Finally, a fixed number of households were randomly selected from the chosen enumeration area. Each primary sampling unit contained province, district, commune, and village identifiers[7]. However, there was little overlap in villages across survey rounds, and thus, analysis mainly took place at the commune level.

Table 2 provides descriptive statistics of the iDE target area and control area[8] prior to the intervention period and the results of a test for baseline balance using CSES data. In addition, the distributions of latrine coverage at baseline are provided in Figure 3. Prior to the intervention, the average latrine coverage rate in the iDE target area is 5.8 percentage points higher compared to areas where no market-based sanitation intervention took place. Since the iDE target area was outperforming the control area prior to the intervention, detecting a treatment effect is more difficult. If an increase in latrine coverage were identified, the superior latrine coverage in the iDE target area prior to the intervention suggests any measured latrine increase is due to the program rather than normal catch-up.

[7] Villages are the smallest administrative unit, followed by communes, which are typically made up of 5-10 villages. Communes are located within districts which are located within provinces. There are 25 provinces in Cambodia; Phnom Penh, the country's capital, is its own province. The map in Appendix B provides further details.

[8] Control provinces are those where neither iDE nor a similar market-based intervention run by WaterSHED was implemented. See Section V for further details.

Figure 3: Comparison of baseline distributions of latrine coverage using CSES data. See Section V for description of the control area.



The CSES serves as a primary means for tracking changes in income, possessions, and employment over time. The survey does not prioritize information on health. In 2004 and 2014, the survey asked households to list the major illness had by each family member in the last month if he or she was sick. This question was used to construct a measure of diarrhea prevalence for children under-5. In 2009, the CSES contained a separate module for children under-5 in which diarrhea occurrence within the last two weeks was more carefully measured. The inconsistency in question type limits the degree to which our calculated diarrhea prevalence outcome can be used to draw strong conclusions when comparing across years.

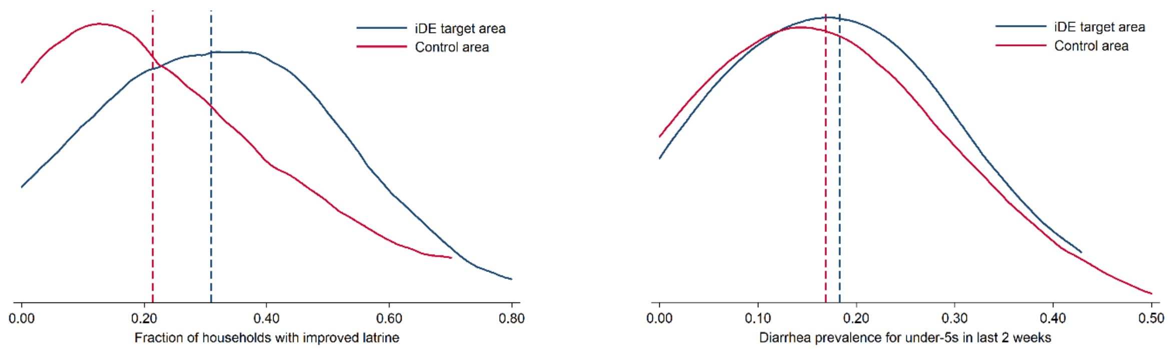
The Demographic Health Survey (DHS) is a USAID-funded initiative wherein detailed health data is collected at the household level across numerous countries to track changing health profiles globally. In Cambodia, DHS data was obtained from 2000, 2005, 2010, and 2014. The enumeration areas were selected using a two-stage stratified sampling approach. In the first stage, the country was stratified by rural-urban areas across 14 individual provinces and 5 pairs of provinces. Enumeration areas were randomly chosen within each stratum, and a fixed number of households were randomly chosen from within each enumeration area. Enumeration areas were geocoded and mapped within contemporaneous districts and provinces using QGIS software. Lower level administrative mapping was not available due to privacy concerns. In contrast to the CSES, the DHS contains very accurate health indicators that are consistently measured over time. However, in creating strata that included pairs of provinces, the survey lacks the same depth of sampling in the smaller provinces as compared to the CSES, which treated each province as its own stratum.

Table 3 provides descriptive statistics of the iDE target area and control area[9] prior to the intervention period and the results of a test for baseline balance for the DHS data. In addition, the distributions of latrine coverage and diarrhea prevalence at baseline are provided in Figure 4a and 4b, respectively. Prior

[9] Control provinces are those where neither iDE nor a similar market-based intervention run by WaterSHED was implemented. See Section V for further details.

to the intervention, the latrine coverage rate in iDE target areas is 9.6 percentage points higher, on average, compared to areas where no market-based sanitation intervention took place. Again, this difference at baseline makes it difficult to capture whether increases in latrine coverage occurred due to iDE's program. From both Table 3 and Figure 4b, one can see that diarrhea prevalence is nearly the same in the iDE target area compared to the control area prior to the intervention.

Figure 4: Comparison of baseline distributions of latrine coverage and diarrhea prevalence using DHS data. See Section V for description of the control area.



(a) Latrine coverage (b) Diarrhea prevalence

Due to the disparate focuses of the three surveys, they vary in terms of the degree to which they accurately measure the different outcomes when compared to each other. Table 1 contains the correlation between the different measures of latrine coverage and diarrhea at baseline and endline at the province level. Since the intervention largely began in 2012, the baseline years are 2012 for iDE latrine count data, 2009 for CSES, and 2010 for DHS. Data was not available from all three sources in the same year prior to the intervention. Endline data was collected in 2014 across all surveys.

Table 1: Correlation in latrine coverage and diarrhea prevalence measures at the province level.

(a) Latrine coverage at baseline.

	iDE	CSES	DHS
iDE	1		
CSES	0.763**	1	
DHS	0.162	0.431**	1

(b) Latrine coverage at endline.

	iDE	CSES	DHS
iDE	1		
CSES	0.969***	1	
DHS	0.870**	0.652***	1

(c) Diarrhea prevalence at baseline

	CSES	DHS
CSES	1	
DHS	0.465**	1

(d) Diarrhea prevalence at endline.

	iDE	CSES	DHS
iDE	1		
CSES	-0.423	1	
DHS	0.0384	-0.0645	1

As already mentioned, since iDE's measure of latrine coverage relies on a village census, it is considered the most accurate measure of that outcome. Looking to Table 1, the CSES data is most correlated to iDE's measure of latrine coverage at both baseline and endline, whereas the DHS data is only similar at endline. In addition, as described previously, DHS has the most accurate measure of diarrhea prevalence and the question formulation stayed constant over time. The results in Table 1 then suggest that the initial measurement of diarrhea in the CSES data was relatively accurate, but that the transition from one question type to another no longer measured the same outcome. Overall, there is high variability in diarrhea prevalence measurements across the different datasets.

Table 2: Descriptive statistics and baseline balance test using 2009 CSES data.

	iDE target area (1)	Control area[a] (2)	(1) vs. (2) t-test
A. Outcome indicators			
Fraction of households with improved latrine	0.248	0.189	0.058*
	(0.017)	(0.031)	(0.035)
Diarrhea prevalence for under-5s in last 2 weeks	0.186	0.139	0.047*
	(0.013)	(0.020)	(0.026)
B. Commune characteristics			
Father's years of schooling	5.878	5.269	0.609***
	(0.079)	(0.148)	(0.165)
Mother's years of schooling	4.605	4.393	0.212
	(0.073)	(0.145)	(0.155)
Consumption per capita per day	933.219	906.627	26.592
	(27.489)	(36.431)	(53.689)
Fraction of households with dirt floors	0.075	0.077	-0.002
	(0.009)	(0.016)	(0.018)
Average number of livestock per household	17.939	19.587	-1.648
	(1.465)	(1.346)	(2.775)
Fraction of households with shared latrine	0.035	0.000	0.035***
	(0.007)	(0.000)	(0.013)
Fraction of households using safe water source	0.528	0.223	0.304***
	(0.029)	(0.038)	(0.057)
Fraction of households treating drinking water	0.561	0.497	0.065*
	(0.018)	(0.034)	(0.038)
Observations[b]	520	128	

* p<0.10, ** p<0.05, *** p<0.01

[a] In the control area, neither iDE's program nor a similar market-based intervention run by WaterSHED was implemented. See Section V for further details.

[b] Observations are at the commune level (weighted average of household characteristics in the commune).

Table 3: Descriptive statistics and baseline balance test using 2010 DHS data.

	iDE target area (1)	Control area[a] (2)	(1) vs. (2) t-test
A. Outcome indicators			
Fraction of households with improved latrine	0.309	0.214	0.096***
	(0.023)	(0.028)	(0.036)
Diarrhea prevalence for under-5s in last 2 weeks	0.183	0.169	0.014
	(0.013)	(0.018)	(0.021)
Weight-for-age SD	-1.471	-1.535	0.064
	(0.047)	(0.048)	(0.068)
Malnourishment prevalence	0.287	0.320	-0.033
	(0.019)	(0.019)	(0.028)
Severe malnourished prevalence	0.077	0.075	0.002
	(0.013)	(0.012)	(0.018)
Weight-for-height SD	-0.745	-0.581	-0.164**
	(0.041)	(0.057)	(0.069)
Wasting prevalence	0.125	0.083	0.042**
	(0.013)	(0.012)	(0.018)
Severe wasting prevalence	0.031	0.016	0.015
	(0.008)	(0.005)	(0.010)
Height-for-age SD	-1.659	-1.949	0.290***
	(0.067)	(0.073)	(0.100)
Stunting prevalence	0.394	0.492	-0.097***
	(0.021)	(0.022)	(0.031)
Severe stunting prevalence	0.147	0.220	-0.073**
	(0.018)	(0.021)	(0.028)
B. District characteristics			
Father's years of schooling	4.381	3.809	0.572**
	(0.165)	(0.214)	(0.266)
Mother's years of schooling	2.653	2.366	0.287
	(0.116)	(0.190)	(0.211)
Wealth index score	-3.879	-4.392	0.512
	(0.454)	(0.708)	(0.805)
Fraction of households with health insurance	0.165	0.214	-0.049
	(0.019)	(0.031)	(0.034)
Fraction of households with dirt floors	0.066	0.066	0.000
	(0.012)	(0.011)	(0.017)
Average number of livestock per household	8.220	9.872	-1.653**
	(0.336)	(0.683)	(0.703)
Fraction of households with shared latrine	0.208	0.172	0.036
	(0.022)	(0.023)	(0.033)
Fraction of households with handwashing facility	0.742	0.626	0.116**
	(0.029)	(0.036)	(0.046)
Fraction of households using safe water source	0.675	0.368	0.308***
	(0.036)	(0.040)	(0.055)
Fraction of households treating drinking water	0.710	0.786	-0.076**
	(0.021)	(0.026)	(0.033)
Observations[b]	65	52	

* p<0.10, ** p<0.05, *** p<0.01

[a] In the control area, neither iDE's program nor a similar market-based intervention run by WaterSHED was implemented. See Section V for further details.

[b] Observations are at the district level (weighted average of household characteristics in the district).

IV.B. Dataset Construction

All of the above-mentioned datasets were combined at the commune, district, and province levels, where possible. Using an additional database that contained the names of all provinces, districts, and communes in the country, the geographic identifiers were harmonized to ensure all observations of the same location across the different survey instruments could be matched together.

Within each dataset, key outcome variables and community characteristics were constructed in accordance with definitions developed by the WHO/UNICEF Joint Monitoring Program (JMP) and standard practices described in the literature (WHO/ UNICEF JMP 2017, Billing et al., 1999). According to the JMP, an improved latrine is a pit latrine with a slab or a flush/pour flush toilet system piped to an alternative location such as a sewer or septic tank (WHO/ UNICEF JMP 2017). Most of the health outcome variables used in the analysis were measured in terms of prevalence using Equation 1.

$$\{health\ outcome\}\ prevalence\ = \frac{number\ of\ children\ under5\ with\ \{health\ outcome\}}{number\ of\ children\ under5\ in\ geographic\ area} \tag{1}$$

In the CSES and DHS, health data for children under-5 were only collected in a subset of households. As such, only those households in which the health measures were collected were used to calculate the total number of children under-5 in the given geographic area. If this were not done, the prevalence of the health conditions would have been lower than the actual condition. A detailed description of the definitions of all outcome variables and community characteristics is further available in the appendix.

Finally, in order to create datasets at various geographic levels, probability-weighted means and totals were calculated. The weighted totals were then used in Equation 1 to calculate the prevalence of each health outcome at the commune, district, or province level. The weighted means formed the basis for the remaining outcome variables and community characteristics. We used the weight variables provided in each of the survey datasets to aggregate the data. The iDE data contained weights to calculate representative averages at the district level. The CSES and DHS data contained weights to calculate representative averages at the rural-urban, by-province strata. Ideally, population data would be used to create new weights to calculate representative averages at other administrative levels. However, data limitations prevented us from taking this approach. As such, the same set of weights was used to calculate probability-weighted means and totals at each geographic level of analysis. It is unclear whether this approach would lead to systematic under or over-estimation of the impacts.

In addition, urban households were excluded from the CSES and DHS data prior to the construction of weighted means and totals. Although urban households could have gained access to latrines through iDE's intervention, the main thrust of their intervention targeted rural areas. Sales agents were only active in rural areas, conducting awareness campaigns and connected customers with latrine producers. Outside research also suggest different population densities between rural and urban areas would lead to differential effects following sanitation interventions (Clasen, et al. 2012; Schmidt 2014). As such, excluding urban households helps ensure we capture the effects that are most probably due to iDE's intervention.

V. Empirical Strategy

iDE's intervention took place in a total of 7 provinces, Banteay Meanchey, Kampong Thom, Kandal, Oddar Meanchey, Prey Veng, Siem Reap, and Svay Rieng. Throughout this report, these provinces are referred to collectively as the iDE target area. At the same time, WaterSHED, another non-profit in Cambodia, began its own sanitation marketing intervention to increase latrine coverage in other regions of the country. Since the Watershed approach is similar to iDE's, it is included as a separate treatment in our analysis to gain a fuller picture of the effects of sanitation marketing. According to their website, WaterSHED's intervention took place in the Battambang, Kampong Cham, Kampong Chhnang, Kampong Speu, Pailin, Pursat, Takeo, and Tbong Khmum provinces.[10] Again, throughout this report, these provinces are referred to collectively as the WaterSHED target area. Excluding Phnom Penh (the capital), eight provinces experienced neither intervention and serve as the control area in our analysis. A map of the country and where each intervention took place is available in Appendix B. The analysis is done in two parts. The first looks at whether any causal impacts can be attributed to iDE's intervention. The second examines the relationship between latrines and health.

V.A. Measuring iDE's Causal Impacts

To determine whether iDE's market-based latrine intervention led to increases in latrine ownership and improvements in health outcomes, we estimate difference-in-difference models specified by Equations 2 and 3. First, Equation 2 is estimated at the commune level using CSES data. Equation 2 will only be used to estimate changes in latrine coverage. Diarrhea prevalence will not be estimated due to inconsistency in measurement methodology as described above.

$$Y_{c,p,t} = \beta_0 + \beta_1 D_p + \beta_2 T_t + \beta_3 (D_p * T_t) + \beta_4 W_p + \beta_5 (W_p * T_t) + \boldsymbol{X_{c,p,t}} + \varepsilon_p \tag{2}$$

$Y_{c,p,t}$ is the outcome for commune c in province p and $\boldsymbol{X_{c,p,t}}$ is a vector of observable characteristics of commune c in province p at time t. D_p and W_p are indicator variables for whether the province was exposed to the iDE program or WaterSHED program, respectively, and T_t is an indicator variable for whether the observation is post-intervention. Standard errors are clustered at the province level. The commune characteristics in $\boldsymbol{X_{c,p,t}}$ include average levels of parental education, average consumption per capita per day, average number of livestock, fraction of households who use improved water sources, fraction of households who treat drinking water, fraction of households who have dirt floors, and fraction of households who share their latrine.

[10] http://watershedasia.org/wp-content/uploads/Report-Sanitation-Coverage-2014-2.pdf

For this estimation, pre-intervention data was collected in 2009 ($T_t = 0$) and post-intervention data was collected in 2014 ($T_t = 1$). Since CSES data was also available from 2004, treatment and control provinces were compared over time to see if they followed parallel trends prior to intervention. Moreover, placebo testing was done in which $T_t = 0$ for data collected in 2004 and $T_t = 1$ for data collected in 2009.

In addition, to measure the effects of iDE's program on health outcomes, Equation 3 is estimated at the district level using DHS data, as follows. Latrine coverage will also be used as an outcome variable in Equation 3 to see whether the DHS and CSES specifications show similar results.

$$Y_{d,p,t} = \beta_0 + \beta_1 D_p + \beta_2 T_t + \beta_3 (D_p * T_t) + \beta_4 W_p + \beta_5 (W_p * T_t) + \boldsymbol{X_{d,p,t}} + \varepsilon_p \quad (3)$$

$Y_{d,p}$ is the outcome for district d in province p and $\boldsymbol{X_{d,p,t}}$ is a vector of observable characteristics of district d in province p at time t. D_p, W_p, and T_t are the same as defined for Equation 2. Standard errors are again clustered at the province level. Here, the district characteristics in $\boldsymbol{X_{d,p,t}}$ include average levels of parental education, average wealth index score, average number of livestock, fraction of households who use improved water sources, fraction of households who treat drinking water, fraction of households who have handwashing facilities, fraction of households who have health insurance, fraction of households who have dirt floors, and fraction of households who share their latrine.

The outcome variables estimated in Equation 3 are latrine coverage; weight-for-age, weight-for-height, and weight-for-height standard deviations; as well as diarrhea, malnourishment, severe malnourishment, wasting, severe wasting, stunting, and severe stunting prevalence. For this estimation, pre- and post-intervention data were collected in 2010 ($T_t = 0$) and in 2014 ($T_t = 1$), respectively. Since DHS data was also obtained from 2000 and 2005, treatment and control provinces were compared prior to iDE's intervention to test for parallel trends and placebo tests were also run in which $T_t = 0$ for data collected in 2005 and $T_t = 1$ for data collected in 2010.

In order to test for the effects of alternative disease transmission pathways, the model in Equation 4 is estimated using DHS data.

$$Y_{d,p,t} = \beta_0 + \beta_1 D_p + \beta_2 T_t + \beta_3 (D_p * T_t) + \beta_4 W_p + \beta_5 (W_p * T_t) + \boldsymbol{X_{d,p,t}} + \boldsymbol{\delta}(D_p * \boldsymbol{X_{d,p,t}}) + \boldsymbol{\eta}(W_p * \boldsymbol{X_{d,p,t}}) + \boldsymbol{\theta}(T_t * \boldsymbol{X_{d,p,t}}) + \boldsymbol{\gamma}(D_p * T_t * \boldsymbol{X_{d,p,t}}) + \varepsilon_p \quad (4)$$

All variables in Equation 4 are the same as those defined in Equation 3. Here, the vector of observable characteristics, $\boldsymbol{X_{d,p,t}}$, is interacted with the iDE province indicator variable (D_p), with the WaterSHED province indicator variable (W_p), with the post-intervention indicator variable (T_t), and with both the iDE province and post-intervention indicator variables ($D_p * T_t$). The observable characteristics, $\boldsymbol{X_{d,p,t}}$, are all demeaned. The outcome variables estimated in Equation 4 will include the diarrhea and physical growth health measures available in the DHS data. The vector of estimated coefficients, $\boldsymbol{\gamma}$, will show whether changes in health in iDE provinces following the intervention are larger or smaller based on the presence of alternative disease transmission pathways.

V.B. Examining Latrine and Health Relationship

To answer whether having a latrine makes a household healthier, the fixed effects model in Equation 5 is estimated using DHS data.

$$Y_{d,p,t} = \boldsymbol{\alpha_{d,p}} + \boldsymbol{\lambda_t} + \beta_1 D_p + \beta_4 W_p + \boldsymbol{X_{d,p,t}} + \varepsilon_p \quad (5)$$

$\boldsymbol{\alpha_{d,p}}$ are the fixed effects for each district d in province p, $\boldsymbol{\lambda_t}$ is a vector of time dummies for years 2010 and 2014, and the other variables are as described previously.

If a causal relationship between increased latrines and better health is difficult to identify, the following set of multivariate regressions is estimated to determine whether latrines and health are associated with one another. First, Equation 6 is estimated using CSES and DHS data from 2014 focusing only on provinces where iDE's intervention took place.

$$Y_{i,hh} = \beta_0 + \beta_1 L_{hh} + \boldsymbol{\varphi}(L_{hh} * \boldsymbol{Z_{hh}}) + \boldsymbol{Z_{hh}} + \boldsymbol{X_{hh}} + \boldsymbol{W_{i,hh}} + \varepsilon_v \quad (6)$$

$Y_{i,hh}$ is the outcome for each child i in household hh, L_{hh} is an indicator for whether the household had a latrine, $\boldsymbol{Z_{hh}}$ is a vector of household-level sanitation-related variables, $\boldsymbol{X_{hh}}$ is a vector of household-level socioeconomic variables, $\boldsymbol{W_{i,hh}}$ is a vector of child-level characteristics, and standard errors are clustered at the village level.

In estimating Equation 6 with CSES data, $\boldsymbol{Z_{hh}}$ includes household sanitation-related characteristics, namely whether the household uses an improved water source, whether drinking water is treated, the number of livestock, whether the flooring is concrete, and whether the latrine is shared. $\boldsymbol{X_{hh}}$ includes household characteristics such as parental education and consumption per capita per day. $\boldsymbol{W_{i,hh}}$ includes the child's age. Where Equation 6 is estimated using DHS data, $\boldsymbol{Z_{hh}}$ includes whether the household uses an improved water source, whether drinking water is treated, whether handwashing facilities are present, the number of livestock, whether the flooring is concrete, and whether the latrine is shared. $\boldsymbol{X_{hh}}$ includes household characteristics such as parental education, the household's wealth index score, and whether the household has health insurance. In summary, the DHS analysis adds handwashing facilities and health insurance to the variables used in the CSES analysis, as well as incorporating a different measure of income.

In addition, we estimated whether the effects of owning a latrine on health outcomes are dependent on the latrine coverage rate in the village where the household is located. In this way, we can determine whether any herd immunity exists where there are more latrines. This answers the question, does having many neighbors with latrines decrease the disease prevalence in the area and make a household healthier even if that household does not have a latrine? This analysis is done by estimating Equation 7.

$$Y_{i,hh,v} = \boldsymbol{\omega}(NL_{hh} * \boldsymbol{TH_v}) + \boldsymbol{\sigma}(L_{hh} * \boldsymbol{TH_v}) + \boldsymbol{X_{hh}} + \boldsymbol{Z_{hh}} + \boldsymbol{W_{i,hh}} + \varepsilon_v \quad (7)$$

$\boldsymbol{NL_{hh}}$ indicates the household does not have a latrine and all other variables are as defined above for Equation 6. CSES data is used to estimate Equation 7. The outcome variable is diarrhea prevalence at the household level. The $\boldsymbol{TH_c}$ vector is a set of thresholds corresponding to whether the village has <30% latrine coverage, between 30% and 70% latrine coverage, or more than 70% latrine coverage. T-tests are done to compare the coefficients in vectors $\boldsymbol{\omega}$ and $\boldsymbol{\sigma}$ to see if having a latrine is associated with any marginal improvement in health, and to see whether having more latrines in the community is associated with improvements in health. DHS data was not used to estimate Equation 7 because there were insufficient villages with above 70% latrine coverage to be able to estimate the equation.

Finally, iDE's own measures of diarrhea and latrine coverage at the village level are used to estimate a relationship between the two using Equations 8 and 9.

$$Y_v = \beta_0 + \beta_1 LC_v + \varepsilon_v \qquad (8)$$

$$Y_v = \beta_0 + \sum_{i=1}^{4} \beta_i * LC_v^i + \boldsymbol{X_c} + \varepsilon_c \qquad (9)$$

For Equation 8, Y_v is the fraction of households where at least one household member had diarrhea in the last month in village v and LC_v is the latrine coverage in village v. This model is estimated locally at intervals of latrine coverage but does not include any observables characteristics of the village and does not allow for a flexible relationship between latrine coverage and diarrhea.

In Equation 9, a flexible functional form for the relationship between latrine coverage and diarrhea is introduced ($\sum_{i=1}^{4} \beta_i * LC_v^i$). In addition, Equation 9 includes a set of observable characteristics of commune c in which village v is located ($\boldsymbol{X_c}$) using variables available in the CSES data. Standard errors are clustered at the commune level in Equation 9.

VI. Results

Having described the methodology, the following section reports the results from each of the model specifications along three main avenues of research: whether changes in latrine coverage and health can be causally linked to iDE's intervention and whether a relationship exists between latrines and health.

VI.A. Did iDE's Intervention Lead to Increases in Latrine Coverage?

Addressing iDE's role in increasing latrine coverage, Table 4 provides the results of estimating Equation 2 using Socio-Economic Survey data aggregated at the commune level. The results show that, in rural areas, iDE's intervention led to increases in latrine coverage compared to other regions of the country. Specifically, after controlling for commune characteristics, the analysis shows that latrine coverage increases by 17.8 percentage points, or a 94% increase, in iDE target areas compared to control areas after their intervention (Column (2) in Table 4). WaterSHED's market-based intervention also led to an increase in latrine coverage of 12.3 percentage points. Similar statistically significant results were found when Equation 2 was estimated using Socio-Economic Survey data at the district level (Appendix, Table C1). Additional analysis was done to measure the effect of the market-based sanitation approach more broadly by pooling the iDE and WaterSHED target areas and can be found in Appendix G.

Table 4: Difference-in-difference analysis of changes to latrine coverage using Socio-Economic Survey data at the commune level.

	Latrine coverage	
	(1)	(2)
iDE target area	0.0584	-0.0192
	(0.070)	(0.047)
After 2012	0.148***	0.0160
	(0.025)	(0.029)
iDE target area X after 2012	0.116**	0.178***
	(0.047)	(0.038)
WaterSHED target area	0.0537	-0.0064
	(0.055)	(0.042)
WaterSHED target area X after 2012	0.0770**	0.123***
	(0.031)	(0.030)
Covariates included[a]	No	Yes
Control mean[b]	0.189	0.189
Observations	1144	1124

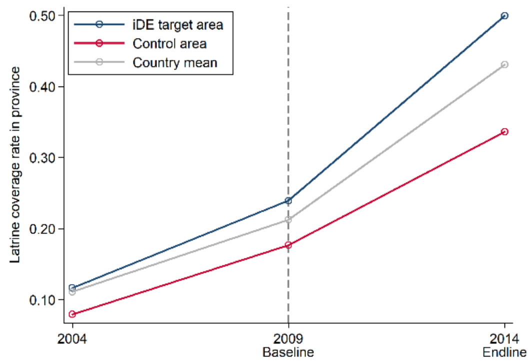
* p<0.10, ** p<0.05, *** p<0.01

[a] Covariates include commune averages for father's and mother's years of schooling, consumption per capita per day, whether water is collected from a clean source, whether drinking water is treated, number of livestock, whether the household flooring is dirt, and whether the latrine is shared.

[b] The reported mean is of the control area prior to the intervention.

We conducted parallel trends analysis to verify that there are no significant differences between the treatment and control areas in the pre-treatment period, thereby validating the use of the difference-in-difference approach. Looking to Figure 5, the treatment and control areas have relatively parallel latrine coverage trends prior to the baseline year. Moreover, the distributions of households with improved latrines look relatively similar at baseline (Section IV, Figure 3). Finally, a placebo test was conducted in which 2009 was treated as the endline measure and 2004 was treated as the baseline, and no significant effects were found due to iDE's intervention (Appendix, Table C2). Taken together, this suggests the chosen methodology was valid and the results are robust.

Figure 5: Parallel trends testing of latrine coverage using CSES data.



Using DHS data aggregated at the district level, the results of estimating Equation 3 are provided in Table 5. As can be seen, Table 5 finds no statistically significant change in latrine coverage in iDE target areas after the intervention, whether or not district characteristics are controlled for.

Table 5: Difference-in-difference analysis of changes to latrine coverage using Demographic Health Survey data at the district level.

	Latrine coverage	
	(1)	(2)
iDE target area	0.0956*	0.0807**
	(0.051)	(0.038)
After 2012	0.141***	0.133***
	(0.023)	(0.024)
iDE target area X after 2012	0.0500	0.0299
	(0.044)	(0.041)
WaterSHED target area	0.127**	0.0767*
	(0.060)	(0.044)
WaterSHED target area X after 2012	-0.00071	-0.0069
	(0.048)	(0.035)
Covariates included[a]	No	Yes
Control mean[b]	0.214	0.214
Observations	337	323

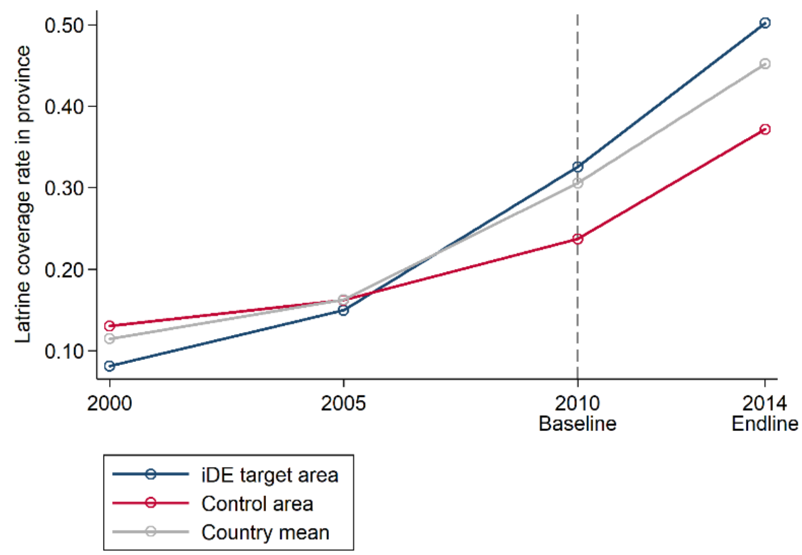
* p<0.10, ** p<0.05, *** p<0.01

[a] Controls include district averages of father's and mother's years of schooling, wealth index score, whether water is collected from a clean source, whether drinking water is treated, whether there are handwashing facilities, whether household has health insurance, number of livestock, whether the household flooring is dirt, and whether the latrine is shared.

[b] The reported mean is of the control area prior to the intervention.

There are several reasons why the DHS results do not replicate the results found using the CSES data. First, as described in Section III, the primary focus of the DHS survey is on taking health measures and its measure of latrine coverage is largely uncorrelated with iDE's more accurate measure of latrine coverage. In addition, the sample size is smaller in the DHS analysis which may result in insufficient statistical power to detect any treatment effects. Moreover, similar robustness analysis was done regarding parallel trends, shown graphically in Figure 6.

Figure 6: Parallel trends testing of latrine coverage using DHS data.



In Figure 6, we do not see parallel trends in latrine coverage in the survey years prior to the 2010 baseline. In addition, in Section IV, Figure 4a, we see different baseline distributions of latrine coverage between treatment and control areas. One explanation could be the way in which DHS defines its sampling strata in comparison to the CSES. As mentioned in Section IV, the DHS combines small provinces together and treats them as one strata, losing sampling depth whereas the CSES and our analysis treat each province separately. This may explain poorer correlation between DHS and iDE measures of latrine coverage and why the CSES results do not replicate when using the DHS data. For all of the above reasons, we believe the CSES results represent the true relationship between iDE's intervention and latrine coverage.

VI.B. Did iDE's Intervention Lead to Improvements in Health?

Looking to iDE's programmatic effects on health, both CSES and DHS contain measures of diarrhea prevalence. However, as described in Section IV, CSES's measure of diarrhea changes across survey rounds, and as a result, cannot be used for the given analysis.[11] Since diarrhea incidence is higher in iDE

[11] When diarrhea is measured with a one-month recall, there is lower variability in diarrhea prevalence than when it is measured using a 2-week recall period. As a result, the measure of diarrhea in 2009 has a larger spread than in either 2004 or 2014.

target areas in 2009, the difference-in-difference estimation will over-estimate iDE's impact whereas using ANCOVA as an alternative will under-estimate the impact. The trends in diarrhea prevalence over time and the results of the two different models using CSES data are included in Appendix D.

Using the more accurate and time-consistent measure of diarrhea in the DHS data, the results of estimating Equation 3 are presented in Table 6. When controlling for district characteristics and holding everything constant, the results show that after IDE's intervention, diarrhea prevalence decreased by 5.8 percentage points, or a 34% decrease, in iDE target areas compared to control areas, at a 5% significance level. A similar effect was not found for WaterSHED target areas. Analysis of the effect of market-based interventions more broadly, pooling the WaterSHED and iDE target areas together, is provided in Appendix G.

Table 6: Difference-in-difference analysis of changes to diarrhea prevalence using Demographic Health Survey data at the district level.

	Diarrhea prevalence	
	(1)	(2)
iDE target area	0.0140	0.00460
	(0.032)	(0.024)
After 2012	-0.0189	-0.0104
	(0.031)	(0.020)
iDE target area X after 2012	-0.0671*	-0.0584**
	(0.036)	(0.027)
WaterSHED target area	-0.0272	-0.0234
	(0.031)	(0.022)
WaterSHED target area X after 2012	0.0164	0.00437
	(0.038)	(0.030)
Covariates included[a]	No	Yes
Control mean[b]	0.169	0.169
Observations	337	323

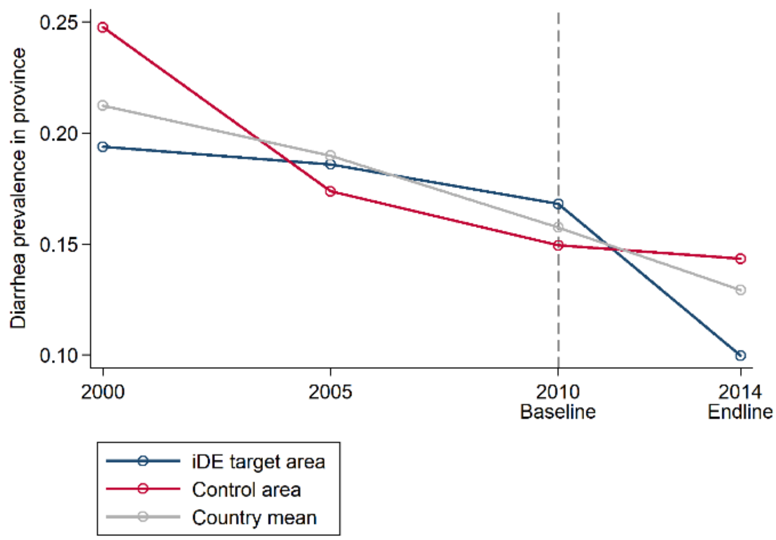
* p<0.10, ** p<0.05, *** p<0.01

[a] Controls include district averages of father's and mother's years of schooling, wealth index score, whether water is collected from a clean source, whether drinking water is treated, whether there are handwashing facilities, whether household has health insurance, number of livestock, whether the household flooring is dirt, and whether the latrine is shared.

[b] The reported mean is of the control area prior to the intervention.

As before, we conduct parallel trends testing to verify that there are no significant differences between the treatment and control areas in the pre-treatment period with regard to diarrhea prevalence. Looking to Figure 7, the diarrhea prevalence trends between control and treatment areas are parallel from 2005-2010, but the trends are not parallel back to 2000. In Section IV, Figure 4b, the baseline distributions of diarrhea prevalence rates are nearly identical for iDE and control areas, suggesting the measured effect is not due to catch-up. Finally, a placebo test was conducted in which 2010 was treated as the endline measure and 2005 the baseline and no significant effects were found due to iDE's intervention (Appendix, Table E1). Taken together, this suggests the difference-in-difference results are robust.

Figure 7: Parallel trends testing of diarrhea prevalence using DHS data.



In addition, we estimated Equation 3 using other physical growth outcome variables, such as weight-for-height, weight-for-age, and height-for-age standard deviations and prevalence of malnutrition, stunting, and wasting, to test the alternative health pathways illustrated in our theory of change. iDE's intervention did not have any effect on these other health outcomes after possible effects were excluded due to failures in parallel trends or pre-treatment balance (see Appendix, Table E2). We hypothesize that we cannot see any improvements in physical growth measures because these health outcomes take a longer time to change as compared to diarrhea. Since iDE's intervention began in 2012 and our endline data is from 2014, it is unsurprising that no effects were found on physical growth measures in two short years.

Furthermore, in line with our theory of change, there is reason to believe that the effect of being in an iDE target area on diarrhea prevalence can be larger or smaller in the presence of other community factors, such as clean drinking water, higher income, etc. Equation 4 was thus estimated to determine whether these other factors influenced iDE's effect on diarrhea prevalence and the physical growth outcomes tested (see Appendix, Table E3). At the 10% significance level, the effect of iDE's intervention on decreasing diarrhea prevalence is larger as more households treat their drinking water and use clean drinking water sources. Furthermore, being in an iDE target area and in a district where more households use clean drinking water sources leads to a statistically significant reduction in malnourishment and severe malnourishment prevalence. Finally, being in an iDE target area and in a district where more households treat their drinking water leads to a statistically significant increase in weight-for-height z-scores and decrease in severe wasting prevalence. Taken together, these results suggest the effect of iDE's latrine intervention on health may be affected by alternative pathways. The results also suggest that measures of physical growth may only change when latrine interventions are in the presence of other disease transmission barriers. However, given the small sample size used to estimate Equation 4, further research with a larger sample size is needed before drawing strong conclusions.

The remaining set of coefficients on other potential mediating factors in Appendix, Table E3 were either insignificant, inconsistent in their effects across different health outcomes, and/or pointed in the wrong direction when compared to the expected direction described in the literature. Thus, there is insufficient evidence to conclude whether these other factors in the disease transmission pathway change the effect iDE's intervention has on the change in diarrhea prevalence, or physical growth measures. The lack of results with regard to the effect of other community characteristics may be due to low statistical power; some of the magnitude of the coefficients were large although not statistically significant. Moreover, these characteristics and health outcomes are all at the district level, preventing us from incorporating household-level variation.

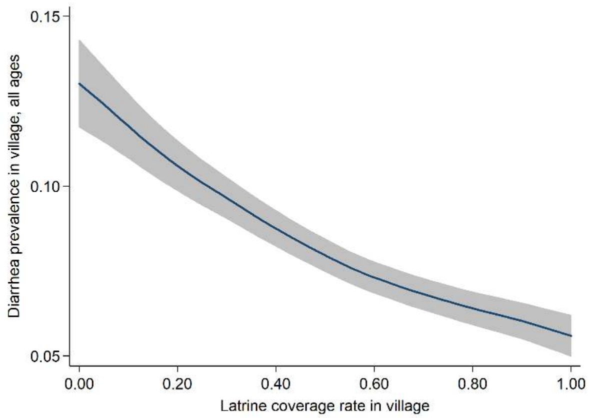
VI.C. What is the Relationship Between Latrine Coverage and Health?

Having found evidence that iDE's intervention led to simultaneous improvements in latrine coverage and diarrhea prevalence, the following set of results attempts to characterize the direct relationship between latrines and health. First, the fixed effects model described in Equation 5 allowed us to determine whether increases in latrine coverage directly caused improvements in health. The results in Appendix F, Table F1 show no statistically significant effects of increasing latrine coverage on the set of health outcomes available in the DHS data. Thus, the data available are insufficient to identify a causal relationship between latrine ownership rates and disease prevalence rates at the district level.

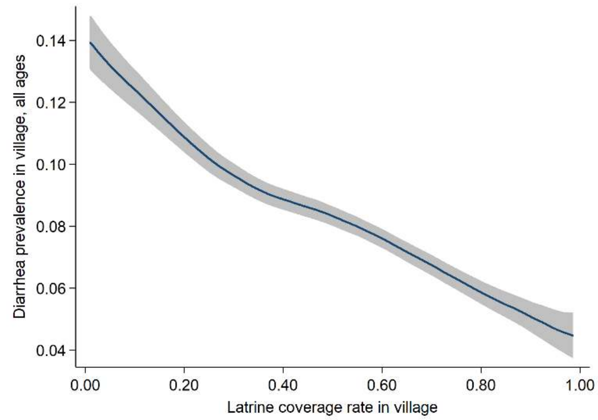
However, the above results are not altogether surprising, as the literature suggests these relationships may only exist at the village or household level (Hammer and Spears, 2013; Pickering et al., 2015; Cameron and Shah, 2017). Given data limitations, it is not possible to estimate a fixed effects model at either the village or household level. Consequently, the following set of results describe whether there is at least an association between increases in latrine coverage and changes in health at the lower levels of geographic aggregation.

iDE's measures are exact rates of latrine coverage and diarrhea at the village level, rather than weighted estimates as in the DHS and CSES data. Therefore, a study of the correlation between the two measures (shown in Figure 8) is likely to present the most accurate picture of the relationship. As stated previously, iDE's measure of diarrhea is a binary indicator of whether someone in the household had diarrhea in the last month.

Figure 8: Graphical representation of the association latrine coverage and diarrhea prevalence in 2016.



(a) Results of local linear regression estimated using Equation 7 (N=837 villages).



(b) Results of overall linear regression estimated using Equation 8, which flexibly controls for latrine coverage and includes commune-level socioeconomic and sanitation characteristics (N=472 villages).

Both with (Figure 8a) and without (Figure 8b) controls, there is a strong association between increases in latrine coverage and decreases in the fraction of households where at least one person has had diarrhea in the last month. The relationship in Figure 8a is statistically significant at the 5% level and suggests an increase in latrine coverage of 10 percentage points is associated with a decrease in diarrhea prevalence of 1.2 percentage points. In Figure 8b, the relationship is significant at the 10% level and suggests a 10-percentage-point increase in latrine coverage is associated with a 5.7 percentage point reduction in diarrhea prevalence. See Appendix F for further details.

Having identified a significant relationship at the village level, we now proceed to examining whether children who live in households with latrines are healthier than those where there is no latrine by estimating Equations 6 and 7. First, the results of estimating Equation 6 using both CSES and DHS data and a range of health outcomes is given in Table 7. The following analysis focuses on the iDE target area in 2014, wherein latrine coverage has already increased due to the sanitation marketing intervention. Since comparison is done across households in a single year, the lack of parallel trends identified previously do not impact the validity of the multivariate regression results.

Table 7: Association between latrine ownership, other sanitation pathways, and health.

	Diarrhea prevalence		Weight-for-age SD	Weight-for-height SD	Height-for-age SD
	(1)	(2)	(3)	(4)	(5)
Has latrine	0.0283	0.0300	-0.592*	0.0161	-1.106*
	(0.070)	(0.245)	(0.356)	(0.266)	(0.592)
Has latrine X has handwashing facilities	-	-0.223	1.091***	1.715***	-0.189
	-	(0.256)	(0.359)	(0.220)	(0.687)
Has latrine X uses improved water source	-0.0172	0.0988	2.271***	1.240***	2.930***
	(0.034)	(0.211)	(0.379)	(0.256)	(0.621)
Has latrine X treats drinking water	0.00858	-0.296	1.525***	0.997***	1.675***
	(0.035)	(0.229)	(0.280)	(0.165)	(0.505)
Has latrine X number of livestock	-0.000	0.00422	-0.200***	-0.144***	-0.224***
	(0.000)	(0.011)	(0.041)	(0.025)	(0.073)
Has latrine X concrete floors	-0.000	0.217	-1.885***	-1.058***	-2.506***
	(0.068)	(0.269)	(0.401)	(0.247)	(0.723)
Has latrine X latrine not shared	-0.0169	-0.00773	0.193	-1.250***	2.273***
	(0.025)	(0.216)	(0.264)	(0.170)	(0.527)
Data source	CSES	DHS	DHS	DHS	DHS
Controls[a]	Yes	Yes	Yes	Yes	Yes
Observations	522	695	439	439	439

* p<0.10, ** p<0.05, *** p<0.01

[a] CSES controls include each of the sanitation factors individually, as well as additional controls for child age, household income, and mother's and father's years of schooling. DHS controls include each of the sanitation factors individually, as well as additional controls for child age, household wealth index score, health insurance, and mother's and father's years of schooling

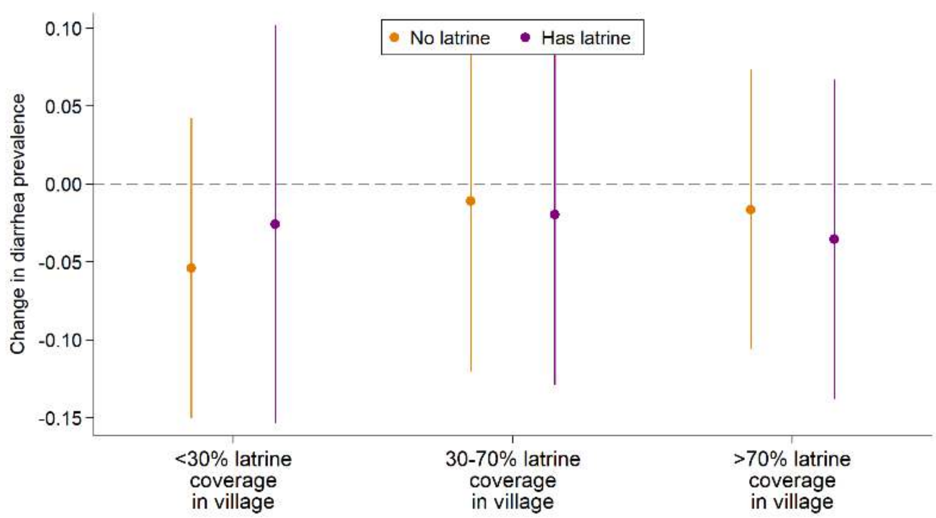
As seen in Columns (1) and (2) of Table 7, owning a latrine has no relationship to diarrhea prevalence at the household level, regardless of whether latrine ownership is combined with other protective practices like using an improved water source, treating drinking water, etc. However, Columns (3) to (5) show relatively consistent effects of the various sanitation pathways on the three measures of physical growth (weight-for-age, weight-for-height, and height-for-age standard deviation).

First, having both a latrine and handwashing facilities is associated with statistically significant increases in weight-for-age and weight-for-height standard deviations. Similarly, having both a latrine and using a clean water source is associated with large, statistically significant increases in all three measures of physical human development. The same can be said for households that have a latrine and treat their drinking water. In addition, children in households with more livestock have statistically significant decreases in their physical growth metrics. These effects all align with the theory of change presented in Section III. Moreover, the effect of combining a latrine and clean drinking water sources or drinking water treatment parallel those found when these factors were to see whether they altered the effect iDE's intervention had on health (Appendix, Table E3). In total, the results of Table 7 suggest having a latrine, combined with a clean water source, treated drinking water, having handwashing facilities, or fewer livestock are associated with better growth outcomes.

The effect of having a shared latrine is inconsistent, and thus, no strong conclusions can be drawn on that regard. Furthermore, having a latrine and concrete floors has the opposite effect of what would be expected, where children's growth measures are worse where there is a latrine and concrete floors. In addition, Table 7 suggests that having a latrine but without the presence of other disease transmission barriers is not associated with any positive improvements in health. This provides moderate evidence in support of interventions that target multiple areas of sanitation, rather than a singular focus on one avenue.

Finally, the literature points to a possible herd immunity benefit to living in a village where a larger fraction of one's neighbors have latrines. For example, a child in a village where nearly all of his or her neighbors have a latrine would be exposed to fewer disease vectors compared to a child in a village where very few households have latrines, regardless of whether that child's household had a latrine. In accordance with the approach used by Andres et al. (2014), thresholds of 30% and 70% village latrine coverage were used. Alternative approaches compare 0% and 100%, but villages meeting these criteria were not present in our dataset. Thus, Equation 9 was estimated to determine whether any herd immunity effects could be estimated using the available data with results presented in Figure 9.

Figure 9: Graphical representation of the coefficients estimated using Equation 9. Comparison of the effect of having a latrine compared to no latrine across villages with different thresholds of larine coverage



As can be seen from Figure 9, there is no statistically significant difference between the diarrhea outcomes of children in households with or without a latrine and compared amongst households in villages with different level of latrine coverage overall. Thus, we found no evidence of herd immunity effects, although the analysis faced data limitations. We were unable to use iDE's measure of latrine coverage, since very few villages were sampled by both iDE and CSES. Instead, we relied on average estimates of village latrine coverage, which face the weaknesses previously described. Moreover, we were unable to use DHS health outcomes since there were insufficient villages with over 70% latrine coverage to be able to estimate the model.

VII. Conclusion

Rural Cambodians suffer from limited access to basic sanitation facilities, which have an important effect on health. To address this sanitation issue, iDE conducted a latrine marketing intervention (SMSU) in seven provinces in Cambodia. This program had three broad aims: (1) to encourage latrine producers to produce improved (pour-flush) latrines; (2) to educate communities on the benefits of improved latrines, and; (3) to mediate sales between households and latrine producers. The SMSU 1.0 program ran from August 2011 to October 2014 facilitating sales of over 141,000 latrines. The Capstone Team aimed to understand the impacts iDE's sanitation intervention had on latrine coverage and health, with a particular focus on changes in diarrhea prevalence for rural Cambodia children under-5.

Based on our analysis using secondary data sources, we found that iDE's sanitation intervention led to improvements in latrine coverage in their targeted area, compared to other areas in Cambodia. Moreover, their intervention led to reductions in diarrhea prevalence for rural Cambodia children under-5. Finally, using iDE's own data, we found that increased latrine coverage at the village level is strongly correlated with decreased rates households where at least one person had diarrhea in the last month. However, we were unable to identify either a causal relationship or an association between latrine ownership at the household level and under-5 diarrhea prevalence when using secondary data sources.

When examining other health outcome variables besides diarrhea, increased latrine coverage combined with some other protective sanitation or household characteristics is strongly associated with three measures of physical growth (weight-for-age, weigh-for-height, and height-for-age standard deviations). Specifically, having a latrine in combination with clean drinking water sources, drinking water treatment, handwashing facilities, and fewer livestock is associated with improvements in a child's physical growth. These effects all align with the theory of change presented in Section III. The effect of having a shared latrine and having a latrine with concrete floors in the household proved inconclusive, however. As well, we found no evidence of herd immunity when investigating heterogeneous effects in villages with different levels of latrine coverage.

Our findings on the impact of iDE's sanitation marketing intervention are generally positive but the study's internal validity is challenged by three limitations. First, each of the surveys were designed to achieve different goals, which affected the construction and accuracy of the variables used in our analysis. Second, these surveys were designed to estimate national and provincial-level statistics. Using the provided provincial weights to construct the commune and district averages using in our analysis may have introduced bias, though it is difficult to determine the direction of the bias a priori. Finally, there is little overlap between the outcomes measured in the primary and secondary data sources and data was not available from all three sources in the same years to capture behavior prior to the intervention in 2012.

In terms of policy implications, our study has shown that market-based interventions, whether conducted by iDE or by WaterSHED, are an effective means of increasing latrine coverage. This finding is robust to the inclusion of community-level characteristics that may affect knowledge of proper sanitation practices and the ability to purchase a latrine. Thus, the primary goal of iDE's and WaterSHED's interventions has been achieved: more people have access to improved sanitation.

These market-based interventions may have resulted in improvements in health, but given the multiple steps between purchase and changes to well-being, the relationship is far from straightforward. First, a latrine must be used consistently for any health effects to be observed. Our analysis also suggests that a latrine may be insufficient to change non-diarrheal health outcomes without additional protective sanitation and hygiene practices. Despite our initial findings, further data is needed to study the effect of latrines in the context of other sanitation practices. Some of these intervening pathways may matter more or less depending on whether one is examining effects at the household-, village-, commune-, or district-level; for example, reaching a particular threshold of latrine coverage might only relevant be at the village-level. Practitioners should think carefully about these relationships when designing their intervention's theory of change and monitoring and evaluation frameworks.

Moreover, WASH practitioners designing surveys to capture household- and community- level characteristics as well as health impacts should think carefully about their methodology. For example, we have seen that conducting censuses are more accurate when doing econometric analysis at lower geographic levels compared to relying on sampling means. In addition, a specific description of what constitutes diarrheal incidence is necessary for household respondents to answer with accuracy. Recall period is also an important factor in receiving reliable responses, as we saw from the differing variability in responses to whether someone had diarrhea when different recall periods were used. Our data suggests two weeks is likely the most accurate.

In addition, further research is needed to firmly establish the causal relationship between increased latrine coverage and improved health impacts. Several of these health impacts may only emerge after many years have passed and may be difficult to identify depending on the geographical level at which the analysis takes place. In addition, more research is needed to evaluate whether latrine interventions should be coupled with other water and sanitation interventions, and whether behavior change campaigns are necessary. Finally, some health impacts, such as diarrhea, may vary according to season, but data often lacks the level of temporal identification needed to measure seasonal variability.

By concentrating on improved latrine coverage, Cambodia can move toward achieving the UN's Sustainable Development Goals (SDG) 3 and 6, thereby ensuring the health of all of its children. iDE's intervention has been shown to be effective in achieving greater household access to latrines and we would encourage them and the Government of Cambodia to work together to ensure the proper level of support, regulation, and financing in order to further expand latrine coverage in the country.

VIII. References

Andres et al. (2014). “Sanitation and externalities: Evidence from early childhood health in rural India.” World Bank Policy Research Working Paper No. 6737.

Billig et al. (1999). “Water and sanitation indictors measurement guide.” Food and Nutrition Technical Assistance Project. Office of Health and Nutrition, Bureau for Global Programs, U.S. Agency for International Development.

Briceño et al. (2017). “Are there synergies from combining hygiene and sanitation promotion campaigns: Evidence from a large-scale cluster-randomized trial in rural Tanzania.” *PLoS ONE*. 12(11): e0186228. https://doi.org/10.1371/journal.pone.0186228.

Cairncross. (1990). “Health impacts in developing countries: New evidence and new prospects.” *Water and Environment Journal*. 4(6): 571-5.

Cairncross and Bartram. (2010). “Hygiene, sanitation, and water: Forgotten foundations of health.” *PLoS Medicine*. 7(11): e1000367. doi:10.1371/journal.pmed/1000367.

Cambodian Ministry of Rural Development, Ministry of Rural Health Care (CMRD). (2010). “National sanitation and hygiene knowledge, attitudes, and practices (KAP) survey.” Accessed 10 Nov 2017. https://camnut.weebly.com/uploads/2/0/3/8/20389289/2010_national_sanitation_and_hygiene_knowledge_attitudes_and_practices_kap_survey_eng_2010.pdf.

Cameron and Shah. (2017). “Scaling up sanitation: Evidence from an RCT in Indonesia.” IZA – Institute of Labor Economics Discussion Papers, No. 10619. http://hdl.handle.net/10419/161242.

Cattaneo et al. (2009). “Housing, health, and happiness.” *American Economic Journal: Economic Policy* 1(1): 75-105.

Chase et al. (2015). “Addressing the behavioural constraints to latrine uptake: Effectiveness of a behaviour-change campaign in rural Cambodia”. *Waterlines.* 34(4): 365-378.

Cheng et al. (2012). “An ecological quantification of the relationships between water, sanitation and infant, child, and maternal mortality.” *Environmental Health*. 11:4. https://www.ehjournal.net/content/11/1/14.

Clasen et al. (2006). “Interventions to improve water quality for preventing diarrhoea. (A Cochrane Review).” In: The Cochrane Library, Issue 3, 2006. Oxford: Update Software.

Clasen et al. (2010). “Interventions to improve disposal of human excreta for preventing diarrhoea (Review).” *Cochrane Database of Systematic Reviews*. Issue 6. No:CD007180. DOI: 10.1002/14651858.CD007180.pub2.

Clasen et al. (2012). “The effect of improved rural sanitation on diarrhoea and helminth infection: design of a cluster-randomized trial in Orissa, India.” *Emerging Themes in Epidemiology*. 9(7). https://doi.org/10.1186/1742-7622-9-7.

Clasen et al. (2014). “Effectiveness of a rural sanitation programme on diarrhea, soil-transmitted helminth infection, and child malnutrition in Odisha, India: a cluster-randomized trial.” *Lancet Global Health.* Issue 2: e645-53.

Dumpert and Perez. (2015). “Going beyond mason training: Enabling, facilitating, and engaging rural sanitation markets for the base of the pyramid.” *Waterlines.* 34(3): 210-226.

Ejemot et al. (2008). “Hand washing for preventing diarrhoea.” Cochrane Database Syst Rev: CD004265

Esrey et al. (1985). “Interventions for the control of diarrhoeal diseases among young children: improving water supplies and excreta disposal facilities.” *Bulletin of the World Health Organization*. 63(4): 757-72.

Esrey et al. (1991). “Effects of improved water supply and sanitation on ascariasis, diarrhea, dracunculiasis, hookworm infection, schistosomiasis, and trachoma.” *Bulletin of the World Health Organization*. 69(5): 609-21.

Exum et al. (2016). “Floors and toilets: Association of floors and sanitation practices with fecal contamination in Peruvian Amazon peri-urban households.” *Environmental Science & Technology*. 50: 7373-7381.

Fewtrell et al. (2005). “Water, sanitation, and hygiene interventions to reduce diarrhea in less developed countries: a systematic review and meta-analysis.” *Lancet Infectious Disease.* 5: 42-52.

Hammer and Spears. (2013). “Village sanitation externalities and children’s human capital: Evidence from a randomized experiment by the Maharashtra government.” World Bank Policy Research. Working Paper No. 6580. https://ssrn.com/abstract=2313535.

Heijnen et al. (2014). “Shared sanitation versus individual household latrines: A systematic review of health outcomes.” *PLoS ONE*. https://doi.org/10.1371/journal.pone.0093300.

Hutton. (2012). “Global costs and benefits of drinking-water supply and sanitation interventions to reach the MDG target and universal coverage”, WHO/HSE/WSH/12.01, Geneva: World Health Organization, [Available at] http://www.who.int/watersanitationhealth/publications/2012/globalcosts.pdf.

Kov et al. (2013). “Growing taller among toilets: Evidence from changes in sanitation and child height in Cambodia, 2005-2010.” Rice Institute.

Kumar and Vollmer. (2012). “Does access to improved sanitation reduce childhood diarrhea in rural India.” *Health Economics*. DOI: 10.1002/hec.2809.

Mosites et al. (2015) "The relationship between livestock ownership and child stunting in three countries in eastern Africa using national survey data." *PLoS ONE.* 10(9):e0136686. doi:10.1371/journal.pone.013668.

Patil et al. (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.

Pickering et al. (2015). "Effect of a community-led sanitation intervention on child diarrhea and child growth in rural Mali: a cluster-randomised controlled trial." *Lancet Global Health*. 3: e701-11.

Prüss et al. (2002). "Estimating the burden of disease from water, sanitation, and hygiene at a global level." *Environmental Health Perspectives.* 110(5): 537-42.

Rah et al. (2015). "Household sanitation and personal hygiene practices are associated with child stunting in rural India: a cross-sectional analysis of surveys." *BMJ Open*. 5: e005180. doi: 10.1136/bmjopen-2014-005180.

Schmidt. (2014). "The elusive effect of water and sanitation on the global burden of disease." *Tropical Medicine and International Health.* 19(5): 522-27.

Smets. (2015). "Water supply and sanitation in Cambodia: turning finance into services for the future." Water and Sanitation Program. World Bank Group

UNICEF. (2017). "Water and sanitation, current status + progress." Accessed 10 Nov 2017. https://data.unicef.org/topic/water-and-sanitation/hygiene/.

Vásquez and Aksan. (2015). "Water, sanitation, and diarrhea incidence among children: evidence from Guatemala." *Water Policy*. 17: 932-45.

Vyas et al. (2016). "Disease externalities and net nutrition: Evidence from changes in sanitation and child height in Cambodia, 2005-2010." *Economics and Human Biology*. 23: 235-42.

Waddington et al. (2009). "Water, sanitation and hygiene interventions to combat childhood diarrhea in developing countries." International Initiative for Impact Evaluation.

WHO/UNICEF JMP. (2017). "Progress on drinking water, sanitation and hygiene: 2017 update and SDG baselines." World Health Organization and the United Nations Children's Fund.

Wolf et al. (2014). "Assessing the impact of drinking water and sanitation on diarrhoeal disease in low- and middle-income settings: a systematic review and meta-regression." *Tropical Medicine and International Health*. 19(8): 928-42.

World Bank (2008a). "Economic impacts of sanitation in southeast Asia: a five-country study conducted in Cambodia, Indonesia, Lao PDR, the Philippines and Vietnam under the Economics of Sanitation Initiative (ESI)." Water and sanitation program research report. Jakarta. World Bank. https://openknowledge.worldbank.org/handle/10986/17262.

World Bank (2008b). "Economic impacts of sanitation in southeast Asia: a four-country study conducted in Cambodia, Indonesia, the Philippines and Vietnam under the Economics of Sanitation Initiative (ESI)." Water and Sanitation Program research report. World Bank. http://documents.worldbank.org/curated/en/246121468231556842/Economic-impacts-of-sanitation-in-Southeast-Asia-a-four-country-study-conducted-in-Cambodia-Indonesia-the-Philippines-and-Vietnam-under-the-Economics-of-Sanitation-Initiative-ESI.

World Food Programme (WFP). (2017). "Cambodia." Accessed 10 Nov 2017. http://www.wfp.org/node/3418.

Yishay et al. (2017). "Microcredit and willingness to pay for environmental quality: Evidence from a randomized-controlled trial of finance for sanitation in rural Cambodia." *Journal of Environmental Economics and Management*. http://dx.doi.org/10.1016/j/jeem.2016.22.004.

VIV. Appendix A: Data Dictionary

Table A1: Definition of key Cambodia Socio-Economic Survey variables used in subsequent analysis.

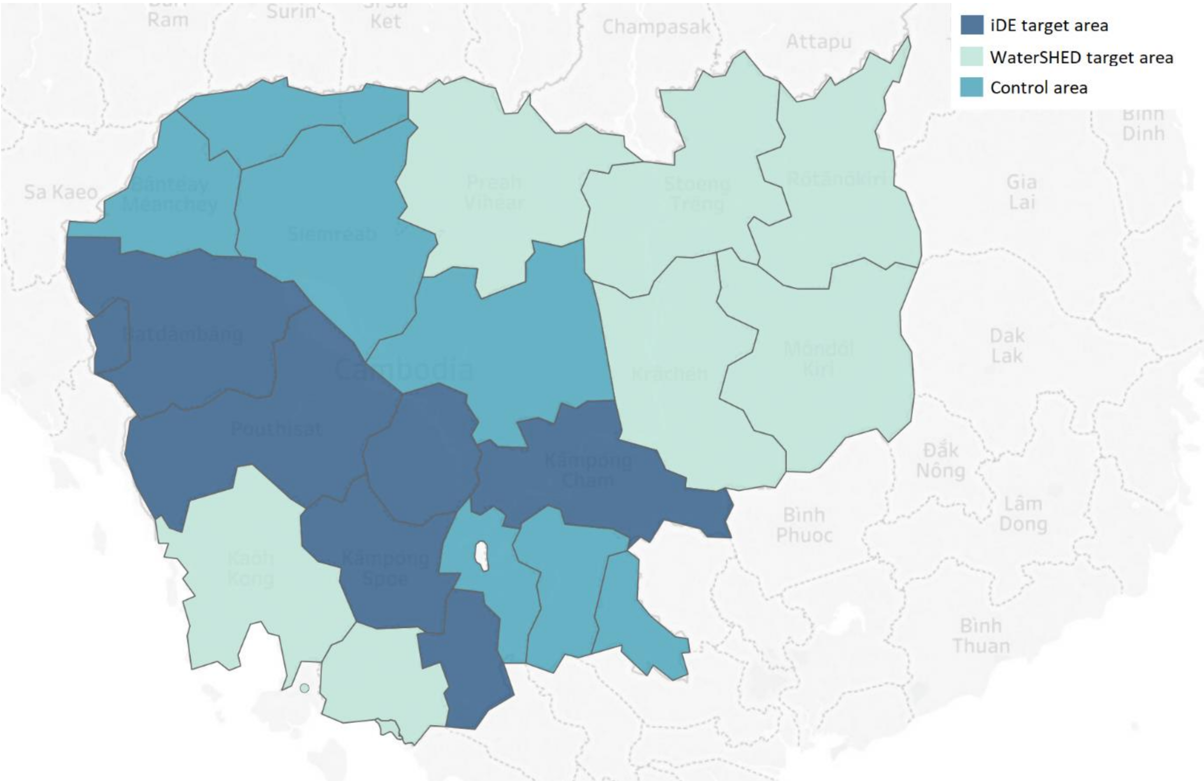
Variable	Description
Latrine coverage	The number of households in the given geographic area that have an improved latrine. As previously defined by the JMP, an improved latrine includes flush/pour flush toilet to piped sewer, septic tank, pit latrine, or a pit latrine with slab.
Diarrhea prevalence	In all cases, the diarrhea prevalence is narrowed to children under, but not including, the age of 5. In 2009, the under-5 module was used to calculate diarrhea prevalence. Children were asked if they had diarrhea in the last 2 weeks. Only children who answered this survey question were included as the baseline for the total number of children under-5 in the given geographic area. In 2014, all household members were asked to identify the most significant illness they had had in the last month. Children under-5 who answered diarrhea were coded as such. Children who answered another condition or did not answer the question at all were assumed not to have diarrhea in the last month.
Father's and mother's years of schooling	The father was identified as the male head of household or male spouse of the head of household. The mother was identified as the female head of household or the female spouse of the health of household. Years of schooling were calculated based on the response to the "highest level of schooling successfully completed." The following set of classifications were made to equate certain levels with years of schooling: - lower secondary school certificate: 11 years - upper secondary school certificate: 12 years - technical/vocational pre-secondary certificate: 12 years - technical/vocational post-secondary certificate: 13 years - college undergraduate: 14 years - bachelor degree: 16 years - masters degree: 18 years - PhD: 22 years
Consumption per capita per day	Households were given a daily diary to list in-kind and in-cash food consumption over a month. The total consumption for the month was calculated and divided by 30 days to approximate consumption per day. That number was then divided by the number of people in the household to approximate consumption per capita per day.
Improved water source	Improved water sources, as defined by JMP, include piped water, boreholes, tubewells, protected dug wells, protected springs, rainwater, and packaged water. Households were only considered as using an improved water source if they used the improved water source during the wet and dry seasons.
Treat drinking water	Households were considered as treating their drinking water if they answered that they always treated water to make it safer for drinking during the last month. If households answered sometimes, they were classified as not treating their drinking water.
Number of livestock	The total number of livestock of any type (cattle, buffaloes, horses, pigs, sheep, goats, chickens, etc.) was calculated for each household.
Concrete floors	Households with concrete floors were all those whose primary flooring material included wooden planks, bamboo strips, cement, brick, stone, wood, vinyl, or ceramic tiles.
Share latrine	Households were said to share their latrine if they answered that the toilet facility they used was a "public toilet/pit latrine or shared with others"

Table A2: Definition of key Demographic Health Survey variables used in subsequent analysis.

Variable	Description
Latrine coverage	The number of households in the given geographic area that have an improved latrine. As previously defined by the JMP, an improved latrine includes flush/pour flush toilet to piped sewer, septic tank, pit latrine, or elsewhere, a ventilated improved pit latrine, or a pit latrine with slab.
Diarrhea prevalence	Not all households were selected to answer questions regarding child health. For those households who were chosen, parents were asked to report whether their child had diarrhea in the last two weeks for every child under, but not including, the age of 5. This was totaled for each geographic level of analysis and divided by the total number of children under-5 in the geographic area who were selected for this survey.
Weight-for-age SD	The weight-for-age standard deviation according to the World Health Organization definition was used.
Malnourishment prevalence	The child's weight-for-age was less than -2 standard deviations from the average. Prevalence was calculated as defined in the main report. Children who were not exposed to this set of survey questions were not include in the total number of children under-5.
Severe malnourishment prevalence	The child's weight-for-age was less than -3 standard deviations from the average. Prevalence was calculated as defined in the main report. Children who were not exposed to this set of survey questions were not include in the total number of children under-5.
Height-for-age SD	The height-for-age standard deviation according to the World Health Organization definition was used.
Wasting prevalence	The child's height-for-age was less than -2 standard deviations from the average. Prevalence was calculated as defined in the main report. Children who were not exposed to this set of survey questions were not include in the total number of children under-5.
Severe wasting prevalence	The child's height-for-age was less than -3 standard deviations from the average. Prevalence was calculated as defined in the main report. Children who were not exposed to this set of survey questions were not include in the total number of children under-5.
Weight-for-height SD	The weight-for-height standard deviation according to the World Health Organization definition was used.
Stunting prevalence	The child's weight-for-height was less than -2 standard deviations from the average. Prevalence was calculated as defined in the main report. Children who were not exposed to this set of survey questions were not include in the total number of children under-5.
Severe stunting prevalence	The child's weight-for-height was less than -3 standard deviations from the average. Prevalence was calculated as defined in the main report. Children who were not exposed to this set of survey questions were not include in the total number of children under-5.
Father's and mother's years of schooling	The father was identified as the male head of household or male spouse of the head of household. The mother was identified as the female head of household or the female spouse of the health of household. The survey reported the exact years of schooling completed by each respondent.
Wealth index score	The wealth index score is a composite measure generated using principal component analysis of a household's ownership of selected assets as defined by the DHS.
Improved water source	Improved water sources, as defined by JMP, include piped water, boreholes, tubewells, protected dug wells, protected springs, rainwater, and packaged water. Households were only considered as using an improved water source if they used the improved water source during the wet and dry seasons.
Treat drinking water	Households were considered as treating their drinking water if they answered that they always or sometimes treated water to make it safer for drinking. This survey did not provide separate answers for always and sometimes, so it was not possible to distinguish the two.
Handwashing facilities	Households were considered as having handwashing facilities if these were observed by the enumerator.
Number of livestock	The total number of livestock of any type (cattle, buffaloes, horses, pigs, sheep, goats, chickens, etc.) was calculated for each household.
Concrete floors	Households with concrete floors were all those whose primary flooring material was not earth, sand or clay.
Latrine shared	Households were asked whether they shared a latrine with other households.
Health insurance	Having health insurance was asked of each child, rather than the household as a whole. For our purposes, households were considered as having health insurance if at least one of their under-5 children had health insurance.

X. Appendix B: Map of Cambodia

Figure B1: Map of Cambodia with provinces identified by whose intervention took place there.



XI. Appendix C: Additional Results on iDE's Effects on Latrine Coverage

Table C1: Difference-in-difference analysis of changes to latrine coverage using Socio-Economic Survey data at the district level.

	Latrine coverage	
	(1)	(2)
iDE target area	0.0792	0.0159
	(0.068)	(0.048)
After 2012	0.155***	0.0143
	(0.023)	(0.031)
iDE target area X after 2012	0.110**	0.174***
	(0.043)	(0.043)
WaterSHED target area	0.104	0.0154
	(0.075)	(0.052)
WaterSHED target area X after 2012	0.0597*	0.133***
	(0.035)	(0.028)
Covariates included[a]	No	Yes
Control mean[b]	0.157	0.157
Observations	322	317

* p<0.10, ** p<0.05, *** p<0.01

[a] Covariates include district averages for father's and mother's years of schooling, consumption per capita per day, whether water is collected from a clean source, whether drinking water is treated, number of livestock, whether the household flooring is dirt, and whether the latrine is shared.

[b] The reported mean is of the control area prior to the intervention.

Table C2: Placebo test of latrine coverage difference-in-difference results using Socio-Economic Survey data at the commune level.

	Latrine coverage	
	(1)	(2)
iDE target area	-0.00410	-0.00616
	(0.037)	(0.022)
After 2004	0.0686	0.0199
	(0.045)	(0.041)
iDE target area X after 2004	0.0625	-0.0102
	(0.059)	(0.045)
Covariates included[a]	No	Yes
Observations	890	881

* p<0.10, ** p<0.05, *** p<0.01

[a] Covariates include commune averages for father's and mother's years of schooling, consumption per capita per day, whether water is collected from a clean source, whether drinking water is treated, number of livestock, whether the household flooring is dirt, and whether the latrine is shared.

XII. Appendix D: Arguments for Excluding CSES Diarrhea Prevalence Causal Analysis

In 2004 and 2014, the survey asked about whether diarrhea occurred using a 1-month recall period whereas in 2009, the recall period was 2 weeks. As can be seen in Figure D1, with a longer recall period, there is less variation in the data, and so any observed effects using difference-in-difference will be picking up measurement error rather than true treatment effects. Thus, the results in Table D1 are an over-estimate of any effect iDE's intervention had on diarrhea prevalence.

Figure D1: Parallel trends testing of diarrhea prevalence using CSES data

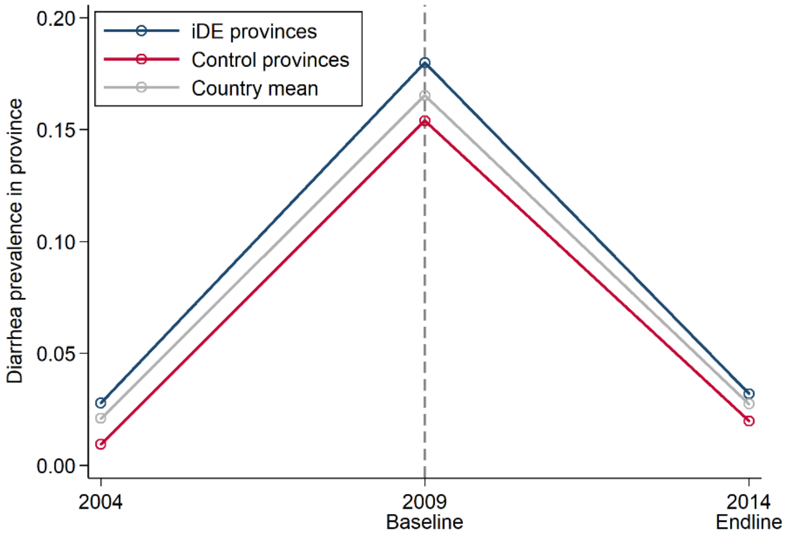


Table D1: Difference-in-difference analysis of changes to diarrhea prevalence using Socio-Economic Survey data at the commune level.

	Diarrhea prevalence	
	(1)	(2)
iDE target area	0.0467**	0.0490**
	(0.018)	(0.020)
After 2012	-0.117***	-0.107***
	(0.012)	(0.014)
iDE target area X after 2012	-0.0324*	-0.0363*
	(0.017)	(0.018)
Covariates included[a]	No	Yes
Control mean[b]	0.139	0.139
Observations	1131	1112

* p<0.10, ** p<0.05, *** p<0.01

[a] Covariates include commune averages for father's and mother's years of schooling, consumption per capita per day, whether water is collected from a clean source, whether drinking water is treated, number of livestock, whether the household flooring is dirt, and whether the latrine is shared.

[b] The reported mean is of the control area prior to the intervention.

Note: regression is done using difference-in-difference and includes an indicator for the WaterSHED target area and the WaterSHED target area interacted with the post period.

To attempt to account for the over-estimate obtained using difference-in-difference, the ANCOVA model described by the following equation was estimated.

$$Y_{c,p,t} = \beta_0 + \beta_1 D_p + \beta_3 W_p + \beta_4 Y_{c,p,t-1} + \boldsymbol{X_{c,p,t-1}} + \varepsilon_p$$

$Y_{c,p,t}$ is the outcome for commune c in province p at time t and $Y_{c,p,t-1}$ is the outcome for commune c in province p at the previous time period $t-1$. D_p and W_p are indicator variables for whether the province was exposed to the iDE program or WaterSHED program, respectively. $\boldsymbol{X_{c,p,t-1}}$ is a vector of observable characteristics of commune c in province p at the previous time period $t-1$. The results of the ANCOVA specification are in Table D2. When commune characteristics are included, there is no statistically significant effect of iDE's intervention on diarrhea prevalence.

Table D2: ANCOVA analysis of changes to diarrhea prevalence using Socio-Economic Survey data at the commune level.

	Diarrhea prevalence	
	(1)	(2)
iDE province	0.0195**	0.0196
	(0.008)	(0.013)
Covariates included[a]	No	Yes
Lagged control mean[b]	0.139	0.139
Observations	254	251

* p<0.10, ** p<0.05, *** p<0.01

[a] Covariates include commune averages for father's and mother's years of schooling, consumption per capita per day, number of livestock, whether the household flooring is dirt, whether water is collected from a clean source, whether drinking water is treated, and whether the latrine is shared, from 2009.

[b] The reported mean is of the control area in 2009.

Note: regression is done using ANCOVA and includes a control for WaterSHED target areas.

XIII. Appendix E: Additional Results on iDE's Effects on Health Outcomes

Table E1: Placebo test of diarrhea prevalence difference-in-difference results using Demographic Health Survey data at the district level.

	Diarrhea prevalence	
	(1)	(2)
iDE target area	-0.00470	0.0250
	(0.043)	(0.028)
After 2005	-0.0178	0.0463
	(0.041)	(0.037)
iDE target area X after 2005	0.0187	-0.0400
	(0.053)	(0.050)
Covariates included[a]	No	Yes
Observations	322	287

* p<0.10, ** p<0.05, *** p<0.01

[a] Controls include district averages of father's and mother's years of schooling, wealth index score, whether water is collected from a clean source, whether drinking water is treated, whether there are handwashing facilities, whether household has health insurance, number of livestock, whether the household flooring is dirt, and whether the latrine is shared.

Table E2: Difference-in-difference analysis of changes to physical growth health outcomes using Demographic Health Survey data at the district level. Regressions in which covariates were excluded were also estimated but are not shown here for conciseness.

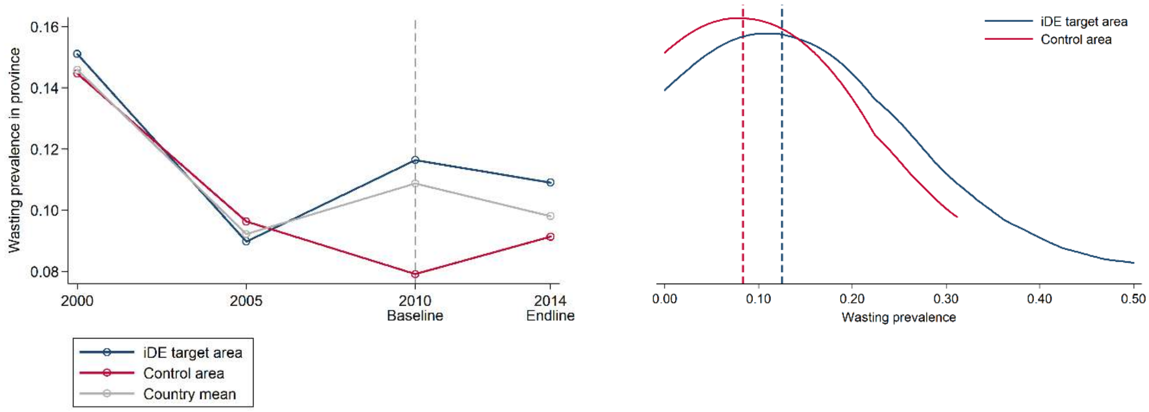
	Weight-for-age SD	Malnourished prevalence	Severe malnourished prevalence	Height-for-age SD	Stunting prevalence	Severe stunting prevalence	Weight-for-height SD	Wasting prevalence	Severe wasting prevalence
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
iDE target area	0.0575	-0.0230	0.00239	0.175*	-0.0683**	-0.0471*	-0.0843	0.0421**	0.0138**
	(0.069)	(0.029)	(0.012)	(0.099)	(0.031)	(0.025)	(0.054)	(0.016)	(0.007)
After 2012	0.163***	-0.0424**	-0.0237	0.268***	-0.0857***	-0.0600**	-0.00778	0.0159	0.00634
	(0.055)	(0.017)	(0.017)	(0.064)	(0.028)	(0.025)	(0.067)	(0.026)	(0.006)
iDE target area X after 2012	-0.0244	-0.0000531	-0.00393	-0.0443	0.0137	-0.000142	0.00920	-0.0518*	-0.00280
	(0.097)	(0.041)	(0.018)	(0.121)	(0.045)	(0.029)	(0.103)	(0.028)	(0.008)
Covariates included[a]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Control mean[b]	-1.535	0.320	0.075	-1.949	0.492	0.220	-0.581	0.083	0.016
Observations	323	323	323	323	323	323	323	323	323

* p<0.10, ** p<0.05, *** p<0.01

[a] Controls include district averages of father's and mother's years of schooling, wealth index score, whether water is collected from a clean source, whether drinking water is treated, whether there are handwashing facilities, whether household has health insurance, number of livestock, whether the household flooring is dirt, and whether the latrine is shared.

[b] The reported mean is of the control provinces prior to the intervention.

Figure E1: Parallel trends testing and comparison of baseline distributions of wasting prevalence using DHS data. Lack of parallel trends invalidates the statistically significant effect of iDE's intervention on wasting prevalence.



(a) Average wasting prevalence over time.

(b) Comparison of wasting prevalence distribution at baseline. The dotted lines represent the average for each group.

Table E3: Exploration of whether iDE's programmatic effect on health outcomes, particularly diarrhea prevalence, varies in the presence of other sanitation or socioeconomic factors.

	Diarrhea prevalence (1)	Weight-for-age SD (2)	Malnourished prevalence (3)	Severe malnourished prevalence (4)	Weight-for-height SD (5)	Wasting prevalence (6)	Severe wasting prevalence (7)	Height-for-age SD (8)	Stunting prevalence (9)	Severe stunting prevalence (10)
iDE target area	0.0130	0.0878	-0.00177	-0.00242	-0.0135	0.0378	0.00223	0.109	-0.0265	-0.00316
	(0.030)	(0.065)	(0.035)	(0.018)	(0.049)	(0.024)	(0.008)	(0.083)	(0.030)	(0.027)
After 2012	-0.0150	0.188**	-0.0133	-0.0225	0.113	-0.00909	-0.00200	0.165***	-0.0423*	-0.0153
	(0.031)	(0.075)	(0.024)	(0.024)	(0.107)	(0.033)	(0.007)	(0.059)	(0.024)	(0.014)
iDE target area X after 2012	-0.0484	-0.0683	-0.0109	0.00623	-0.0796	-0.0306	0.0167	0.00582	-0.0247	-0.0392
	(0.036)	(0.130)	(0.043)	(0.026)	(0.156)	(0.040)	(0.010)	(0.096)	(0.042)	(0.024)
iDE target area X after 2012 X household wealth index score	0.0248***	-0.0390	0.00981	-0.00469	-0.00536	-0.0149	-0.00654	-0.0912	0.0293*	-0.00429
	(0.007)	(0.040)	(0.011)	(0.007)	(0.026)	(0.009)	(0.005)	(0.058)	(0.016)	(0.011)
iDE target area X after 2012 X father's years of schools	-0.00818	-0.0271	-0.0370	0.0289	0.0932	0.0149	0.00405	-0.111	0.0249	-0.0114
	(0.033)	(0.151)	(0.063)	(0.037)	(0.109)	(0.040)	(0.016)	(0.251)	(0.064)	(0.048)
iDE target area X after 2012 X mother's years of schools	-0.00340	0.0698	0.0350	0.0164	-0.182	0.0570	0.0406*	0.357	-0.119	-0.0865
	(0.040)	(0.245)	(0.072)	(0.036)	(0.181)	(0.053)	(0.021)	(0.337)	(0.072)	(0.069)
iDE target area X after 2012 X clean drinking water source	-0.309*	0.401	-0.405*	-0.221*	0.211	0.0523	0.00900	0.484	-0.351	-0.159
	(0.153)	(0.552)	(0.222)	(0.117)	(0.497)	(0.131)	(0.052)	(0.729)	(0.227)	(0.221)
iDE target area X after 2012 X treat drinking water	-0.220*	0.407	-0.167	-0.0403	1.329*	-0.388	-0.232***	-1.421	-0.0450	0.183
	(0.128)	(0.680)	(0.371)	(0.129)	(0.704)	(0.234)	(0.066)	(0.968)	(0.291)	(0.239)
iDE target area X after 2012 X handwashing facilities	-0.0885	-0.167	0.246*	0.000303	-0.782	-0.0536	-0.0958*	0.530	-0.139	0.0202
	(0.212)	(0.555)	(0.140)	(0.125)	(0.464)	(0.174)	(0.049)	(0.708)	(0.167)	(0.112)
iDE target area X after 2012 X health insurance	0.0527	-0.762	-0.114	-0.0138	-1.238	0.664***	0.150*	0.553	-0.0713	-0.205
	(0.282)	(0.764)	(0.208)	(0.167)	(0.862)	(0.197)	(0.079)	(1.110)	(0.329)	(0.131)
iDE target area X after 2012 X concrete floors	-0.0385	0.0584	-0.178	0.0658	-0.186	0.0254	0.0773**	0.309	-0.109	-0.0380
	(0.073)	(0.210)	(0.125)	(0.054)	(0.234)	(0.063)	(0.028)	(0.242)	(0.093)	(0.080)
iDE target area X after 2012 X latrine not shared	0.0695	-0.398	0.237	-0.0251	0.0367	-0.0314	-0.0623*	-0.749**	-0.0132	0.0360
	(0.086)	(0.323)	(0.152)	(0.086)	(0.354)	(0.087)	(0.032)	(0.337)	(0.168)	(0.092)
iDE target area X after 2012 X number of livestock	0.0107	-0.00246	-0.00833	-0.00483	0.0710***	-0.0159	-0.00674***	-0.0842*	0.0193	-0.0113
	(0.011)	(0.031)	(0.008)	(0.011)	(0.019)	(0.012)	(0.002)	(0.046)	(0.013)	(0.013)
Covariates included[a]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	323	323	323	323	323	323	323	323	323	323

* p<0.10, ** p<0.05, *** p<0.01

Note: Analysis is done at the district level. Regression includes WaterSHED target areas, WaterSHED target areas interacted with the post period, and WaterSHED target areas interacted with the post period and each mediating factor. Regression also includes each mediating factor and each mediating factor interacted with the post period.

XIV. Appendix F: Additional Results on Relationship Between Latrines and Health

Table F1: Results of fixed effects specification using district-level Demographic Health Survey data attempting to establish causal relationship latrine coverage and health.

	Diarrhea prevalence (1)	Weight-for-age SD (2)	Malnourished prevalence (3)	Severe malnourished prevalence (4)	Height-for-age SD (5)	Stunting prevalence (6)	Severe stunting prevalence (7)	Weight-for-height SD (8)	Wasting prevalence (9)	Severe wasting prevalence (10)
Latrine coverage rate	-0.0632	0.0190	-0.0570	0.0615	0.436	-0.0632	0.0337	-0.382	0.116	0.0451
	(0.092)	(0.201)	(0.073)	(0.037)	(0.318)	(0.117)	(0.077)	(0.279)	(0.086)	(0.033)
Covariates included[a]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	323	323	323	323	323	323	323	323	323	323

* p<0.10, ** p<0.05, *** p<0.01

[a] Covariates include district averages of father's and mother's years of schooling, wealth index score, whether water is collected from a clean source, whether drinking water is treated, whether there are handwashing facilities, whether household has health insurance, number of livestock, whether the household flooring is dirt, and whether the latrine is shared.

Table F2: Association between village latrine coverage and fraction of households in the village experiencing diarrhea, using iDE data.

	Fraction of households with at least 1 member with diarrhea in last month (1)	(2)
Village latrine coverage rate	-0.119**	-0.568*
	(0.048)	(0.317)
Covariates included[a]	No	Yes
Observations	837	472
R-squared	0.13	0.23

* p<0.10, ** p<0.05, *** p<0.01

[a] Covariates include commune averages for father's and mother's years of schooling, consumption per capita per day, whether water is collected from a clean source, whether drinking water is treated, number of livestock, whether the household flooring is dirt, and whether the latrine is shared.

XV. Appendix G: Analysis of Effects of Market-Based Interventions

iDE's and WaterSHED's interventions are both market-based sanitation programs and are similar in nature. As such, the following set of results treats the two sets of provinces as a single treatment area that was exposed to a market-based sanitation intervention beginning in 2012. In this analysis, there are now 15 treatment provinces and 9 control provinces. Tables G1 and G2 provide descriptive statistics of the outcomes and community characteristics used in the analysis, as well as the results of a statistical comparison of means.

Table G1: Measuring the effect of market-based sanitation interventions, descriptive statistics and baseline balance test using 2009 Socio-Economic Survey data.

	Had market-based intervention (1)	Control area (2)	(1) vs. (2) t-test
A. Outcome indicators			
Fraction of households with improved latrine	0.245	0.189	0.056*
	(0.011)	(0.031)	(0.032)
B. Commune Characteristics			
Father's years of schooling	5.794	5.269	0.524***
	(0.057)	(0.148)	(0.158)
Mother's years of schooling	4.617	4.393	0.224
	(0.051)	(0.145)	(0.146)
Consumption per capita per day	925.169	906.627	18.542
	(16.162)	(36.431)	(44.089)
Fraction of households with dirt floors	0.081	0.077	0.004
	(0.006)	(0.016)	(0.016)
Average number of livestock per household	18.064	19.587	-1.523
	(0.925)	(1.346)	(2.466)
Fraction of households with shared latrine	0.034	0.000	0.034***
	(0.005)	(0.000)	(0.014)
Fraction of households using safe water source	0.423	0.223	0.199***
	(0.019)	(0.038)	(0.051)
Fraction of households treating drinking water	0.556	0.497	0.059*
	(0.012)	(0.034)	(0.034)
Observations[a]	758	128	

* p<0.10, ** p<0.05, *** p<0.01

[a] Observations are at the commune level (weighted average of household characteristics in the commune).

Table G2: Measuring the effect of market-based sanitation interventions, descriptive statistics and baseline balance test using 2010 Demographic Health Survey data.

	Had market-based intervention (1)	Control area (2)	(1) vs. (2) t-test
A. Outcome indicators			
Fraction of households with improved latrine	0.325	0.214	0.111***
	(0.018)	(0.028)	(0.034)
Diarrhea prevalence for under-5s in last 2 weeks	0.163	0.169	-0.006
	(0.010)	(0.018)	(0.019)
Weight-for-age SD	-1.488	-1.535	0.047
	(0.033)	(0.048)	(0.061)
Malnourishment prevalence	0.293	0.320	-0.027
	(0.015)	(0.019)	(0.027)
Severe malnourished prevalence	0.070	0.075	-0.005
	(0.008)	(0.012)	(0.015)
Weight-for-height SD	-0.767	-0.581	-0.186**
	(0.033)	(0.057)	(0.064)
Wasting prevalence	0.124	0.083	0.041**
	(0.010)	(0.012)	(0.017)
Severe wasting prevalence	0.029	0.016	0.013
	(0.005)	(0.005)	(0.009)
Height-for-age SD	-1.670	-1.949	0.280***
	(0.048)	(0.073)	(0.090)
Stunting prevalence	0.392	0.492	-0.099***
	(0.017)	(0.022)	(0.031)
Severe stunting prevalence	0.139	0.220	-0.081**
	(0.012)	(0.021)	(0.023)
B. District Characteristics			
Father's years of schooling	4.606	3.809	0797**
	(0.113)	(0.214)	(0.226)
Mother's years of schooling	2.809	2.366	0.443**
	(0.084)	(0.190)	(0.179)
Wealth index score	-3.705	-4.392	0.687
	(0.311)	(0.708)	(0.666)
Fraction of households with health insurance	0.157	0.214	-0.057*
	(0.014)	(0.031)	(0.030)
Fraction of households with dirt floors	0.075	0.066	0.009
	(0.008)	(0.011)	(0.015)
Average number of livestock per household	8.535	9.872	-1.338**
	(0.293)	(0.683)	(0.635)
Fraction of households with shared latrine	0.240	0.172	0.068**
	(0.017)	(0.023)	(0.032)
Fraction of households with handwashing facility	0.674	0.626	0.048
	(0.029)	(0.036)	(0.052)
Fraction of households using safe water source	0.572	0.368	0.206***
	(0.027)	(0.040)	(0.050)
Fraction of households treating drinking water	0.726	0.786	-0.060**
	(0.014)	(0.026)	(0.028)
Observations[a]	125	52	

* p<0.10, ** p<0.05, *** p<0.01

[a] Observations are at the district level (weighted average of household characteristics in the district).

In order to determine whether market-based interventions lead to changes in latrine coverage, we estimated the difference-in-difference models described by Equations 2 and 3. Table G3 contains the latrine coverage impacts, estimated using both the CSES and DHS data. As can be seen in column (2) of Table G4, exposure to market-based sanitation intervention led to a 15 percentage point increase in latrine coverage compared to control areas.

We conducted parallel trends analysis to verify that there are no significant differences between the treatment and control areas in the pre-treatment period, thereby validating the use of the difference-in-difference approach. Looking to Figure G1a, the treatment and control areas have relatively parallel latrine coverage trends prior to the baseline year. Moreover, the distributions of households with improved latrines look relatively similar at baseline (Figure G1b). Finally, a placebo test was conducted in which 2009 was treated as the endline measure and 2004 was treated as the baseline, and no significant effects were found due to market-based interventions (Table G4). Taken together, this suggests the methodology was valid and the CSES results are to be believed.

The DHS findings (columns (3)-(4) of Table G3) do not replicate the results found with the CSES data. In explanation, we do not see parallel trends in latrine coverage in Figure G2a in the survey years prior to the 2010 baseline. In addition, in Figure G2b, we see different distributions of latrine coverage between treatment and control areas. The data-driven reasons for the discrepancy between CSES and DHS results can be found in the main report.

Table G3: Measuring the effect of market-based sanitation interventions, difference-in-difference analysis of changes in latrine coverage using commune-level Socio-Economic Survey data and district-level Demographic Health Survey data.

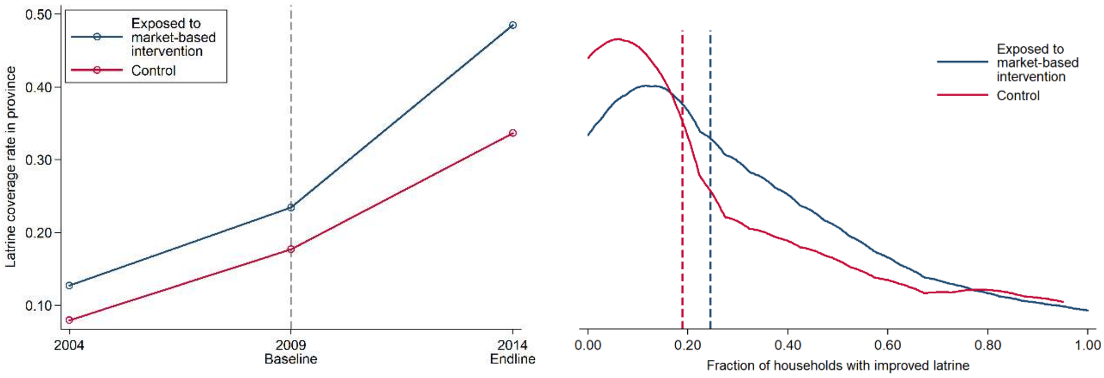
	Latrine coverage			
	(1)	(2)	(3)	(4)
Exposed to market-based intervention	0.0560	-0.0139	0.111**	0.0783**
	(0.051)	(0.040)	(0.049)	(0.034)
After intervention	0.148***	0.0158	0.141***	0.133***
	(0.025)	(0.028)	(0.023)	(0.025)
Exposed to market-based intervention X after intervention	0.0957***	0.150***	0.0257	0.0131
	(0.033)	(0.031)	(0.037)	(0.031)
Data source	CSES	CSES	DHS	DHS
Administrative level of analysis	commune	commune	district	district
Covariates included[a]	No	Yes	No	Yes
Control mean[b]	0.245	0.245	0.325	0.325
Observations	1144	1124	337	323

* p<0.10, ** p<0.05, *** p<0.01

[a] CSES covariates include commune averages of father's and mother's years of schooling, consumption per capita per day, whether water is collected from a clean source, whether drinking water is treated, number of livestock, whether the household flooring is dirt, and whether the latrine is shared. DHS covariates include district averages of father's and mother's years of schooling, wealth index score, whether water is collected from a clean source, whether drinking water is treated, whether household has handwashing facilities, whether household has health insurance, number of livestock, whether the household flooring is dirt, and whether the latrine is shared.

[b] The reported mean is of the control area prior to the intervention.

Figure G1: Parallel trends testing and comparison of baseline distributions of latrine coverage using CSES data for determining effect of market-based intervention.



(a) Average latrine coverage over time.

(b) Comparison of latrine coverage distribution at baseline. The dotted lines represent the average for each group.

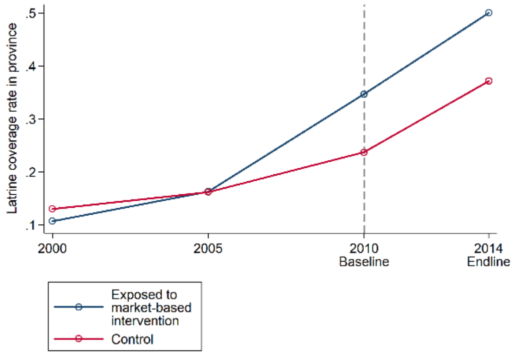
Table G4: Measuring the effect of market-based sanitation interventions, placebo test of latrine coverage difference-in-difference results using Socio-Economic Survey data at the commune level.

	Latrine coverage	
	(1)	(2)
Exposed to market-based intervention	0.0123 (0.033)	0.00592 (0.023)
After 2005	-0.0178 (0.041)	0.0200 (0.041)
Exposed to market-based intervention X after 2005	0.0437 (0.050)	-0.0171 (0.041)
Covariates included[a]	No	Yes
Observations	890	881

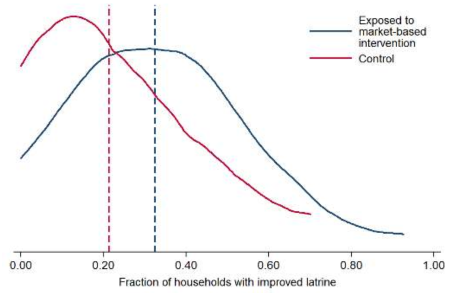
* p<0.10, ** p<0.05, *** p<0.01

[a]Covariates include commune averages of father's and mother's years of schooling, wealth index score, whether water is collected from a clean source, whether drinking water is treated, number of livestock, whether the household flooring is dirt, and whether the latrine is shared.

Figure G2: Parallel trends testing and comparison of baseline distributions of latrine coverage using DHS data for determining effect of market-based intervention.



(a) Average latrine coverage over time.



(b) Comparison of latrine coverage distribution at baseline. The dotted lines represent the average for each group.

The difference-in-difference model was also used to estimate whether market-based sanitation interventions led to changes in health. The results of those estimations can be found in Tables G5 and G6 and suggest that market-based interventions have no effect on health. The statistically significant results for decreased wasting prevalence are not valid once trends over time and baseline distributions are examined (see Figure G3). The results on diarrhea prevalence in Table G5 contradict those found in the main report, where iDE's intervention was shown to decrease diarrhea prevalence. One explanation for this difference may be that little is known about how WaterSHED's intervention was implemented, and as such, their behavior change communication may not be as effective in causing households to actually use their latrines.

Table G5: Measuring the effect of market-based sanitation interventions, difference-in-difference analysis of changes to diarrhea prevalence using Demographic Health Survey data at the district level.

	Diarrhea prevalence	
	(1)	(2)
Exposed to market-based intervention	-0.00578	-0.00903
	(0.030)	(0.021)
After intervention	-0.0189	-0.0117
	(0.030)	(0.020)
Exposed to market-based intervention X after intervention	-0.0271	-0.0294
	(0.036)	(0.027)
Covariates included[a]	No	Yes
Control mean[b]	0.163	0.163
Observations	337	323

* p<0.10, ** p<0.05, *** p<0.01

[a] Controls include district averages of father's and mother's years of schooling, wealth index score, whether water is collected from a clean source, whether drinking water is treated, whether household has handwashing facilities, whether household has health insurance, number of livestock, whether the household flooring is dirt, and whether the latrine is shared.

[b] The reported mean is of the control provinces prior to the intervention.

Table G6: Measuring the effect of market-based sanitation interventions, difference-in-difference analysis of changes to physical growth health outcomes using Demographic Health Survey data at the district level. Regressions in which covariates were excluded were also estimated but are not shown here for conciseness.

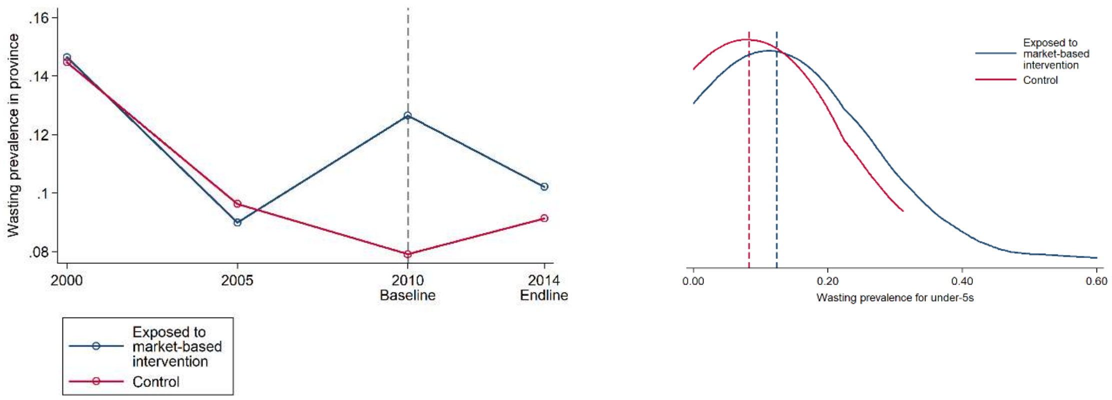
	Weight-for-age SD	Malnourished prevalence	Severe malnourished prevalence	Weight-for-height SD	Wasting prevalence	Severe wasting prevalence	Height-for-age SD	Stunting prevalence	Severe stunting prevalence
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Exposed to market-based intervention	0.0438	-0.0160	-0.000980	-0.0988*	0.0444***	0.0151**	0.164*	-0.0731**	-0.0416
	(0.057)	(0.025)	(0.012)	(0.051)	(0.012)	(0.006)	(0.080)	(0.027)	(0.025)
After intervention	0.163***	-0.0423**	-0.0238	-0.00866	0.0160	0.00655	0.268***	-0.0862***	-0.0600**
	(0.054)	(0.017)	(0.016)	(0.066)	(0.025)	(0.005)	(0.064)	(0.028)	(0.025)
Exposed to market-based intervention X after intervention	-0.0321	0.00460	-0.00652	0.0241	-0.0505*	-0.00923	-0.0737	0.0245	0.00595
	(0.073)	(0.029)	(0.017)	(0.075)	(0.025)	(0.008)	(0.096)	(0.038)	(0.028)
Covariates[a]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Control mean[b]	-1.488	0.293	0.070	-0.767	0.124	0.029	-1.670	0.392	0.139
Observations	323	323	323	323	323	323	323	323	323

* p<0.10, ** p<0.05, *** p<0.01

[a] Controls include district averages of father's and mother's years of schooling, wealth index score, whether water is collected from a clean source, whether drinking water is treated, whether household has handwashing facilities, whether household has health insurance, number of livestock, whether the household flooring is dirt, and whether the latrine is shared.

[b] The reported mean is of the control provinces prior to the intervention.

Figure G3: Parallel trends testing and comparison of baseline distributions of wasting prevalence using DHS data for determining effect of market-based intervention. Lack of parallel trends invalidates the statistically significant effect of iDE's intervention on wasting prevalence.



(a) Average wasting prevalence over time.

(b) Comparison of wasting prevalence distribution at baseline. The dotted lines represent the average for each group.