Institute for Economic Studies, Keio University

Keio-IES Discussion Paper Series

Do Toilets Save Young Children's Lives? Evidence from Cambodia

山田浩之、Tien Manh Vu

2024 年 3 月 28 日
DP2024-007
https://ies.keio.ac.jp/publications/23665/

Institute for Economic Studies, Keio University
2-15-45 Mita, Minato-ku, Tokyo 108-8345, Japan
ies-office@adst.keio.ac.jp
28 March, 2024

Do Toilets Save Young Children’s Lives? Evidence from Cambodia

山田浩之、Tien Manh Vu

IES Keio DP2024-007
2024 年 3 月 28 日
JEL Classification: I15, J13, O18, O53
キーワード: Toilet, Sanitation, Under-five mortality, Census, Cambodia

【要旨】

There are 3.5 billion people living without safe toilets worldwide, who often have no choice but to use unreliable, inadequate toilets or engage in open defecation. The negative impacts of lacking appropriate toilets are particularly noticeable in infants and children, including diseases, stunted growth, chronic malnutrition, and in some cases, death. However, evidence showing these negative effects based on nationally representative data is lacking and thus a thorough analysis is needed. Therefore, in this study, we explore how the extent of toilet coverage is associated with under-five mortality in Cambodia. We use the censuses conducted in 2008 and 2019 in Cambodia to create village-level panel data that include information on the extent of toilet coverage and what kinds of toilets are used. We use the constructed village-level panel data to perform an instrumental variable regression analysis aimed at elucidating the association of toilet coverage with under-five mortality at the village level. We find that increased toilet coverage in a given village is associated with reduced under-five mortality in that village. Increased coverage of improved toilets in particular is associated with lower under-five mortality, suggesting that cleaner toilets save young children’s lives. This finding is useful to policymakers in developing countries facing challenges regarding the widespread use of toilets.

山田浩之
慶應義塾大学経済学部
東京都港区三田 2-15-45
hyamada@econ.keio.ac.jp

Tien Manh Vu
中央大学国際経営学部
東京都八王子市東中野 742-1
mvu264@g.chuo-u.ac.jp

謝辞：This work was supported by JSPS (Japan Society for the Promotion of Science) KAKENHI Grant Numbers 21K01455, 22K18534, and Keio University Academic Development Funds.

Do Toilets Save Young Children's Lives? Evidence from Cambodia

Hiroyuki Yamada* and Tien Manh Vu**

* (Corresponding author) Faculty of Economics, Keio University
2–15–45 Mita, Minato–ku, Tokyo 108–8345, Japan
Tel (Direct): +81-(0)3-5427-1271
E-mail: hyamada@econ.keio.ac.jp
https://orcid.org/0000-0003-3709-1914

** Faculty of Global Management, Chuo University
742-1 Higashinakano, Hachioji, Tokyo 192-0393, Japan
Email: mvu264@g.chuo-u.ac.jp
https://orcid.org/0000-0002-2251-0795

Acknowledgements: This work was supported by JSPS (Japan Society for the Promotion of Science) KAKENHI Grant Numbers 21K01455, 22K18534, and Keio University Academic Development Funds.

Conflict of interest: The authors declare no conflicts of interest.

Abstract

There are 3.5 billion people living without safe toilets worldwide, who often have no choice but to use unreliable, inadequate toilets or engage in open defecation. The negative impacts of lacking appropriate toilets are particularly noticeable in infants and children, including diseases, stunted growth, chronic malnutrition, and in some cases, death. However, evidence showing these negative effects based on nationally representative data is lacking and thus a thorough analysis is needed. Therefore, in this study, we explore how the extent of toilet coverage is associated with under-five mortality in Cambodia. We use the censuses conducted in 2008 and 2019 in Cambodia to create village-level panel data that include information on the extent of toilet coverage and what kinds of toilets are used. We use the constructed village-level panel data to perform an instrumental variable regression analysis aimed at elucidating the association of toilet coverage with under-five mortality at the village level. We find that increased toilet coverage in a given village is associated with reduced under-five mortality in that village. Increased coverage of improved toilets in particular is associated with lower under-five mortality, suggesting that cleaner toilets save young children's lives. This finding is useful to policymakers in developing countries facing challenges regarding the widespread use of toilets.

Keywords: Toilet, Sanitation, Under-five mortality, Census, Cambodia
JEL Classification Codes: I15, J13, O18, O53

1. Introduction

Access to toilets may seem like a basic right, and one taken for granted by many. This issue is strongly related to the sixth goal of the Sustainable Development Goals: "Ensure access to water and sanitation for all." However, there are 3.5 billion people living without safe toilets in the world, who often have no choice but to use unreliable, inadequate toilets or engage in open defecation.[1] Without proper toilets and sanitation systems, untreated human feces can pose a serious threat to public health by contaminating drinking water sources, rivers, and food crops, thereby spreading deadly diseases among the wider population. The negative impacts of not having appropriate toilets are particularly noticeable in infants and children, including diseases, stunted growth, chronic malnutrition, and in some cases, death.

Policymakers in developing countries have been aware of the importance of toilets for a long time. A famous example is China's "toilet revolution." Although no studies have focused on the origin of the toilet revolution in China, the term "toilet revolution" was proposed first by United Nations International Children's Emergency Fund (UNICEF) in 1997, when UNICEF and the National Patriotic Health Campaign Committee worked together to promote toilet retrofitting in China (Cheng et al., 2018). Another notable example is the Governments of India's Total Sanitation Campaign (TSC), a low-subsidy regime that aims to generate household involvement and demand responsiveness for the building of individual household toilets in households below the poverty line (Barnard, et al. 2013). A further example is Community-Led Total Sanitation (CLTS), a program implemented in more than 60 countries in Asia, Africa, Latin America, the Pacific, and the Middle East to address the sanitation burden (da Silva Wells and Sijbesma, 2012). CLTS aims to create demand for sanitation by facilitating graphic, shame-inducing community discussions of the negative health consequences of existing sanitation practices, rather than through the more traditional approach of providing sanitation hardware or subsidies (Cameron et al., 2019).

Researchers are keen to investigate the impacts of toilet utilization on various outcomes. Indeed, previous studies have investigated the impacts of having adequate toilets on outcomes such as self-rated health (Chen et al. 2022), diarrheal disease (Anteneh and Kumie 2010, Semba et al. 2011), labor supply (Wan and Shen 2022), and groundwater quality (Graham 2013, Ndoziya et al. 2019). Spears (2020) claims that open defecation in India might account for Indian children's lower height compared with children in Africa. Using a randomized control trial in the context of CLTS in Laos, Cameron et al. (2021) show that improved sanitation produced positive health spillovers, in that increased village sanitation coverage decreased the probability of childhood stunting.

However, when it comes to the impact of toilets on infant and under-five mortality, the literature remains scarce. Semba et al. (2011) investigate the association between the presence of an improved toilet in the household and under-five mortality in Indonesia. Agha (2000) also investigates the correlation between infant mortality and toilet sanitation. The most closely related paper to ours is Spears (2012), which investigates the causal impact of rural sanitation on infant mortality in the context of TSC in India.

[1] https://www.un.org/en/observances/toilet-day/background (las accessed on March 5, 2024)

However, as we explain in detail in a later section, the empirical identification approach of Spears (2012) (difference-in-differences) differs from our own.

To fill this gap in the literature, we investigate the impact of toilets on under-five mortality in Cambodia by using village-level panel data. It is no exaggeration to say that under-five mortality rate is one of the most important indicators reflecting the sanitary environment of a country. In addition, the survival of children is extremely important for the subsequent accumulation of human capital and, ultimately, for the development of the country. Cambodia is an ideal country to test the hypothesis that toilets have an impact on under-five mortality because the country experienced rapid economic growth between 2008 and 2019 (above 6% annually on average), which led to a dramatic increase in toilet coverage from 35% to 83%.

Our study makes four contributions to the literature concerning the relationship between toilets and under-five mortality. First, we utilize two censuses conducted in 2008 and 2019 in order to construct the village-level panel dataset, which is quite unique in research on the impact of toilets. Because they are censuses, not a sampled survey dataset, they are free from the sample selection problem.

Second, as mentioned above, few studies have investigated the impact of toilets on under-five mortality. This is partly because there are few reliable data on under-five mortality in developing countries. For example, respondents might be inclined to underreport the death of infants and young children because it is a sensitive issue. Utilizing two censuses in Cambodia enables us to avoid this potential problem because they include a death module asking about family members who passed away in the 12 months prior to "census night"—the specific point in time that the collected census information refers to—as well as the causes of death.

Third, we use a novel instrumental variable for the variables regarding toilets to obtain a causal relationship between the availability of toilets and the under-five mortality in the village, if any. Many previous studies merely performed correlation analyses between toilets and health outcomes. Although some studies utilized policy interventions to draw a causal relationship between toilets and health outcomes, none studied the impact of toilets on under-five mortality.

Fourth, in our analysis, both the positive and negative externalities of toilet issues within a village are fully taken into account because we utilize the full censuses and use "village" as the sample unit (i.e., village-level panel data). It is widely known that a household having a toilet can have a positive impact on neighboring households, which is indeed a positive externality. Conversely, open defecation because a household does not have a toilet can lead to negative health outcomes for neighboring households (i.e., a negative externality). Using survey data from a small sample, we cannot tell anything about such positive and negative externalities. However, our village-level panel data using full censuses allows us to consider such externalities.[2]

The rest of this paper is organized as follows. The next section explains the background information and country context of Cambodia. Section 3 explores the data. Section 4 shows the empirical models and the estimation results. In Section 5, we discuss

[2] Cameron et al. (2021) show positive health externality (or spillover) within a village of improved sanitation, but they use a sample of 160 villages. Furthermore, their outcome variable is childhood stunting, not under-five mortality. Geruso and Spears (2018) also examine the importance of neighborhoods by finding that Muslim children are substantially more likely than Hindu children to survive to their first birthday.

the obtained results as well as the policy implications. Section 6 presents the conclusions.

2. Background and context

Cambodia, officially the Kingdom of Cambodia, is a country in Mainland Southeast Asia. The population was about 16.2 million and the population density was 91.8 people per square kilometer of land area in 2019. The population is relatively young, with persons aged 0–14 years accounting for about 30% of the population in 2019. This is largely because the early 1970s were a time of escalating civil war, and in the late 70s during the Khmer Rouge period, a large number of people were killed in Cambodia (Ministry of Planning of Cambodia 2020).

Between the 2008 and 2019 censuses, Cambodia recorded a high mean annual GDP growth rate of more than 6%, even taking into consideration the zero growth rate in 2009 due to the global financial crisis. Hence, Cambodia experienced both economic and social changes during this period.

According to the General Population Census of Cambodia (GPCC) of 2019, which is one of our main data sources for the analysis below, about 83% of households in Cambodia had at least one toilet at home, a great improvement compared with just 35% of households in 2008. Despite this progress, the sanitation issue in Cambodia remains a concern. Indeed, the under-five mortality rate in Cambodia (26.80 per 1,000 live births) in 2019 was much higher than that of neighboring countries such as Vietnam (21.10), Thailand (9.00), and Malaysia (7.90); only Laos had a higher rate (45.80). Furthermore, the mortality rate attributed to unsafe water, unsafe sanitation, and lack of hygiene in Cambodia (17.10 per 100,000 population) is also higher compared with neighboring countries (6.90 in Vietnam, 11.80 in Thailand, and 14.40 in Malaysia).[3]

Although this is a very worrisome situation, solid evidence on the impact of the toilet installation on mortality remains scarce. Therefore, in this paper, we focus in particular on the impact of having a toilet at home on under-five mortality.

3. Data

3.1. Outcome variables

We use the following data sources for the empirical analyses. The first data source is the GPCC of 2008 and 2019, which are full censuses that contain information identifying each village and include a death module asking about household family members who passed away in the 12 months prior to the census night. The death module enables us to count how many children under the age of 5 years passed away during those 12 months. We obtained the under-five mortality rate for each village in 2008 and 2019 by dividing the number of under-five deaths by the number of under-five children who were alive on the census night, multiplied by 100. We did not include under-five deaths unrelated to sanitation, such as those caused by road accidents, land mines, and drowning. Note that although we use the term "under-five mortality" in this paper, we mean under-five mortality in the 12 months prior to the census night. This definition differs from the

[3] The data for population and GDP are taken from World Development Indicators (https://databank.worldbank.org/source/world-development-indicators) (last accessed on March 10, 2024).

one used by the World Health Organization (WHO): "the probability that a child born in a specific year or period will die before reaching the age of 5 years, subject to the age-specific mortality rates of that period".[4] Hence, our computed under-five mortality is smaller than that of the WHO.

We calculate the under-five mortality rate of boys and girls in a similar manner. These three constructed variables—total under-five mortality rate ($Rdeathunder5$), under-five mortality of boys ($Rdeathunder5B$), and under-five mortality of girls ($Rdeathunder5G$)—are our outcome variables.

We omit the top 0.7% of villages in terms of the highest $Rdeathunder5$ (over 15 in value) to overcome the outlier problem.

3.2. Variables of interest

The variables for toilets at the village level also come from the GPCC, in which each household is asked about the type of toilet in the house. First, the percentage of households having no toilet ($Rnotoilet$) in each village is obtained by dividing the number of households that have no toilet at home by the total number of households in the village. Second, the percentage of households having toilets regardless of type ($Rtoilet$) in each village is obtained by dividing the number of households that have a toilet at home by the total number of households in the village.[5] Third, we consider the type of a toilet. We define (a) a pour flush (or flush) toilet connected to sewage and (b) a pour flush (or flush) toilet with a septic tank as an "improved" toilet. The percentage of households having an improved toilet ($Rimprovedtoilet$) in each village is obtained by dividing the number of households that have an improved toilet at home by the total number of households in the village. Then, we define the percentage of households having a "non-improved" toilet ($Rnotimprovedtoilet$) as the number of households having a toilet other than (a) and (b) above divided by the total number of households in the village. A typical non-improved toilet is a pit latrine with or without a slab, or a pour flush (or flush) toilet connected to something other than sewage or a septic tank.

3.3. Instrumental variable

As we discuss in detail in a later section, we suspect that toilet conditions in villages may be endogenous. To overcome this endogeneity, we apply an instrumental variable (IV) approach. The instrumental variable used is information on how steep the slope within the village is. The steepness of the land within the village may be associated with the difficulty or suitability for the construction of a given type of toilet. For instance, in the construction of a toilet with a septic tank, it is generally recommended to slope the

[4] See the following page for details of the WHO's definition of the mortality rate of children under the age of 5:
https://www.who.int/data/nutrition/nlis/info/under-five-mortality-rate#:~:text=How%20is%20it%20defined%3F,mortality%20rates%20of%20that%20period (last accessed on March 15, 2024)

[5] Some households did not answer the question regarding toilets. Hence, $Rnotoilet$ and $Rtoilet$ do not necessarily add up to 1 in all villages. This is the why we use both variables as variables of interest in this paper.

pipe 1/4 inch per foot (1/8 inch per foot minimum) toward the tank.[6] The slope is not directly related to the under-five mortality in the village. To measure the slope of a village, we use the following procedure. First, we have the location of each village as a point on a map in the shape file obtained from the National Institute of Cambodia. Figure 1 shows the location distribution of villages. Unfortunately, the shape file of the villages is not available, so we cannot identify the exact topographic information of each village. Hence, we assume that each village is small, with a radius of 300 m. By observing the point location of each village, this approximation would not be a bad one, although it apparently contains some noise. Second, by overlapping the circle of each village with the topographic information from the Digital Elevation Model created by the Shuttle Radar Topography Mission in 2000 and the ASTER global digital elevation model in 2011, we find the highest and the lowest elevations and take the difference of the two. This procedure is possible because the resolution is 1 arc-second (approximately 30 m) for the two models above. We interpret the difference in elevation between the highest and the lowest elevations as the slope of the village. We use the constructed slope variable using the data from 2000 as an IV for the toilet variable in 2008 and that using the data from 2011 as an IV for the toilet variable in 2019.

3.4. Control variables

All control variables in the main specifications below were obtained from Commune Village Database (CVD) for 2007 and 2018 collected by the National Institute of Statistics (NIS) of Cambodia, except for population in a village, which was obtained from the GPCC 2008 and GPCC 2019. In each year since 2006, NIS has collected various information at the village and/or commune level as the CVD. Village leaders and key informants of the government development commune committee collect the data after training, and monitoring and evaluation are carried out by the Ministry of Planning and the Ministry of Interior. We use the CVD for 2007 and 2018, one year before the census years (i.e., 2008 and 2019, respectively). We construct a variable for the percentage of households whose head is female and the percentage of households whose main water source is from a well, pond, or river, respectively. The illiteracy rates of people aged 15–17, 18–24, 25–35, and 36–45 years are obtained for the male and female populations separately. For housing conditions, we construct a variable for the percentage of households having zinc or fibro roof houses with electricity and those having tiled roof houses with electricity. Variables for the percentage of households using chemical fertilizers and those using pesticides are also included because agriculture remains the main economic activity in many village economies. In addition, the average number of vehicles per household and the distance to the nearest health center from the village are also included.[7]

---

[6] For instance, see the following website https://extension.missouri.edu/publications/eq401 (last accessed on March 15, 2024)

[7] For population and the distance to the nearest health center, we apply the inverse hyperbolic sine transformation. The inverse hyperbolic sine transformation of x is denoted by the following equation: $ln\left(x + \sqrt{x^2 + 1}\right)$.

3.5. Descriptive statistics

Table 1 shows the descriptive statistics of the variables to be used in the analyses below. Total sample size is 24,178 and covers 12,089 of the roughly 14,000 villages in Cambodia, indicating that we have well-balanced panel data. The discrepancy between the number of villages in the sample and the actual number of villages in Cambodia is due mainly to three reasons. First, there are some villages with missing data and/or with outliers (e.g., the number of under-five children and deaths), which we dropped from our sample. Second, some villages in 2008 were no longer villages in 2019 and vice versa, and thus we could not make village-level panel data for these villages. Third, between the two censuses conducted in 2008 and 2019, Kampong Cham Province was split into Kampong Cham and Tboung Khmum provinces. It was difficult to match the villages in Kampong Cham Province in the 2008 census with those in Tboung Khmum Province in the 2019 census because their province, district, and commune codes had changed. Thus, we did not include these villages (about 870 villages) in our sample.

The mean total under-five mortality rate was 0.66%, and the under-five mortality rate for girls (0.79%) was higher than that for boys (0.54%). In a typical village, 53% of households had toilets and 46% of households did not. In addition, 46% of households had improved toilets, while 7% had non-improved toilets. The mean $slope$ was 12.8 m, with a large standard deviation. The descriptive statistics of the controls are shown in Table 1.

4. Empirical model and results

4.1 Empirical model

We apply a standard two-period panel specification with fixed effects. The sample unit is "village." Specifically, we estimate the following empirical model:

$$mortality_{vt} = \beta_0 + \beta_1 toilet_{vt} + control'_{vt}\beta_2 + \varphi_v + \pi_t + \varepsilon_{vt}, \quad (1)$$

where $mortality_{vt}$ is one of the three outcome variables, namely, $Rdeathunder5$, $Rdeathunder5B$, or $Rdeathunder5G$ in village v in year t (=2008 or 2019). $toilet_{vt}$ is our main variable of interest. We have $Rnotoilet$, $Rtoilet$, $Rimprovedtoilet$, and $Rnotimprovedtoilet$ in village v in year t (=2008 or 2019) as a candidate of $toilet_{vt}$. In the specification, we test these variables one by one. $control_{vt}$ is a vector of control variables in village v in year t (=2008 or 2019). φ_v is a village fixed effect and π_t is a time dummy taking 1 if the year is 2019 and 0 otherwise. ε_{vt} is the error term.

However, as briefly mentioned in the previous section, the toilet variable may be endogenous. That is, villages with a certain percentage of households with toilets or no toilets may be non-randomly located in a certain area of Cambodia. Or, simply, the allocation of toilets may be correlated with unobservable cofounders in the error term in the main specification. Hence, a simple estimation of (1) using ordinary least squares (OLS) may yield biased estimates.

To overcome this endogeneity problem, we apply the instrumental variable approach. As explained in the previous section, the IV is the slope of the village. We assume that the slope within a village is related to the difficulty or suitability for the construction of a given type of toilet. Specifically, the slope could ensure the smooth flow of running water within the village or help to flush excrement into septic tanks. However, the slope variable does not imply a direct cause of higher or lower under-five mortality. Note that in computing under-five mortality, we consider only non-accident causes. Hence, the slope variable is an ideal candidate for the toilet variables. Specifically, we have the following first-stage specification:

$$toilet_{vt} = \gamma_0 + \gamma_1 slope_{vt} + control'_{vt}\gamma_2 + \omega_v + \tau_t + \mu_{vt}, (2)$$

where $slope_{vt}$ is the slope variable in village v in year t (=2008 or 2019), as explained in the previous section. ω_v is a village fixed effect and τ_t is a time dummy taking 1 if the year is 2019 and 0 otherwise. μ_{vt} is the error term. Armed with this first-stage specification (2), we estimate the second-stage specification (1) in which the toilet variable is replaced with the fitted value from (2). Because the sample unit is "village," not an individual within the village, robust standard errors are used.

4.2 Results

4.2.1. OLS results

The upper half of Table 2 presents the estimation results of the availability of toilets, using specification (1) with OLS (i.e., without IV). It shows the estimation results using total under-five mortality, under-five mortality for boys, and under-five mortality for girls, both with and without control variables.

The percentage of households without toilets in a village is associated with higher total and boys' under-five mortality in the village but not with girls' under-five mortality (Columns (1), (2), and (3), respectively). The coefficient of $Rnotoilet$ is positive and statistically significant for total and boys' under-five mortality. The coefficient of $Rnotoilet$ for girls' under-five mortality is positive but is not statistically significant. These results hold even if the control variables are included (Columns (4), (5), and (6)).

The percentage of households with toilets in a village is associated with lower total and boys' under-five mortality in the village but not with girls' under-five mortality (Columns (7), (8), and (9), respectively). The coefficient of $Rtoilet$ is negative and statistically significant for total and boys' under-five mortality. The coefficient of $Rtoilet$ for girls' under-five mortality is negative but is not statistically significant. These results hold even if the control variables are included (Columns (10), (11), and (12)).

In the bottom half of Table 2, we decompose $Rtoilet$ into two variables: $Rimprovedtoilet$ and $Rnotimprovedtoilet$. The former is the percentage of households having improved toilets in a village, while the latter is the percentage of households having non-improved toilets in a village; both were defined in the previous section.

We find that the percentage of households with improved toilets in a village is negatively associated with total, boys', and girls' under-five mortality in the village (Columns (13)-(18)). For the under-five mortality of boys, the coefficient of $Rimprovedtoilet$ is negative but statistically insignificant in Column (14). However, when the control variables are included, the coefficient of $Rimprovedtoilet$ becomes negative and statistically significant. Total and girls' under-five mortality are negative and statistically significant both with and without the control variables.

However, the percentage of households with non-improved toilets in a village is not associated with under-five mortality. The coefficient of $Rnotimprovedtoilet$ does not have any statistical significance except Column (24), which implies that the percentage of households having non-improved toilets in a village is positively associated with under-five mortality for girls, implying non-improved toilets are harmful to the survival of young girls.

The results in Table 2 reveal that the popularization of toilets in a given village is negatively associated with lower under-five mortality in that village, but that association is entirely driven by the popularization of improved toilets in the village.

4.2.2. IV results

The results obtained from the OLS are insightful, but they may suffer from the endogeneity of the toilet variables, our main variables of our interest, as discussed earlier. Hence, we perform an IV estimation for specification (1), and its first stage is (2) in 4.1. The IV is the slope variable constructed from the highest and lowest altitude within a village. The slope could ensure the smooth flow of running within the village or help to flush excrement into septic tanks. Or, conversely, the slope might make it difficult to construct a toilet. Whether a steeper slope helps or hinders the construction of toilets is an empirical question; however, the slope itself is not directly related to under-five mortality. The latter condition is an exclusion restriction of the IV that must be satisfied. Ultimately, there is no established test that definitively investigates the validity of instruments in the context of exclusion restrictions (Kiviet, 2020). However, this condition seems to hold in our context because we do not include under-five deaths unrelated to sanitation, such as those caused by road accidents, land mines, and drowning, when computing under-five mortality. Therefore, the slope variable of each village is a good candidate IV for the toilet variables.

Table 3 presents the estimation results of the availability of toilets using the IV estimations. From Panel B of Columns (1)–(6), we find that the IV is strongly and negatively statistically correlated with the percentage of households without toilets in a village. The coefficient of $slope$ is negative and statistically significant at the 1% level. The value of the F test of the excluded instrument is large enough (52.97 for Columns (1)–(3) and 93.19 for Columns (4)–(6)), rejecting the hypothesis that the IV is weak. Therefore, we can proceed to the interpretation of the second-stage results for Columns (1)–(6).

From Panel A of Columns (1)–(3), we find that the percentage of households without toilets in a village has a negative impact on the survival of young children, in terms of total, boys', and girls' under-five mortality, but its impacts are not statistically significant. However, when the control variables are appropriately included in Columns (4)–(6), the percentage of households without toilets in a village has a negative impact on total, boys', and girls' under-five mortality, and the estimated impacts are statistically significant at the 5% level for total and boys' under-five mortality and at the 10% level for girls' under-five mortality. The estimated impact is large. For instance, a 10-percentage point increase in households without toilets in a village increases the total under-five mortality rate by 0.221 percentage points. The negative impact on under-five mortality is larger for boys than for girls.

Conversely, from Columns (7)–(12), we find that the percentage of households having toilets in a village reduces total, boys' and girls' under-five mortality. From Panel

B of Columns (7)–(12), the IV is strongly and positively statistically correlated with the percentage of households having toilets in a village. The value of the F test of the excluded instrument is large enough to reject the hypothesis that the IV is weak. From panel A, although the coefficient of $Rtoilet$ is not statistically significant without the control variables in Columns (7)–(9), it becomes statistically significant with the control variables in Columns (10)–(12). Again, the estimated impact is large. For instance, a 10-percentage point increase in households having toilets in a village decreases the total under-five mortality rate by 0.218 percentage points. The positive impact of having a toilet in a village on under-five mortality is larger for boys than for girls.

In Table 4, we decompose $Rtoilet$ into two variables, $Rimprovedtoilet$ and $Rnotimprovedtoilet$, and conduct IV estimations separately for $Rimprovedtoilet$ and $Rnotimprovedtoilet$.

Columns (1)–(6) of Table 4 present the IV estimation results using $Rimprovedtoilet$ as the main variable of interest. From Panel A, the IV is strongly and positively statistically correlated with the percentage of households having improved toilets in a village. The value of the F test of the excluded instrument is large enough to reject the hypothesis that the IV is weak. Panel B shows the second-stage estimation results. Although the coefficient of $Rtoilet$ is not statistically significant without the control variables in Columns (7)–(9), it becomes statistically significant with the control variables in Columns (10)–(12). The estimated magnitude of the impacts is quite similar to that obtained using $Rtoilet$ in Table 3. For instance, a 10-percentage point increase in the percentage of households with improved toilets in a village decreases the total under-five mortality rate by 0.227 percentage points. The positive impact of having improved toilets in a village on under-five mortality is larger for boys than for girls.

Meanwhile, $slope$ is not a valid IV for the percentage of households with non-improved toilets. From Panel B of Columns (7)–(12), $slope$ is not statistically correlated with $Rnotimprovedtoilet$ at all and, from the value of the F test of the excluded instrument (1.12 in Columns (7)–(9) and 0.22 in Columns (10)–(12)), we cannot reject the hypothesis that the IV is weak. Therefore, we cannot make any causal inference regarding the impact of the percentage of households having non-improved toilets in a village on under-five mortality. In other words, we would need a good IV for non-improved toilets, but we leave this issue for a future study.

Our IV estimation results provide similar implications as the ones obtained from the OLS results. Combining the results in Tables 3 and 4, the popularization of toilets in a given village reduces total, boys', and girls' under-five mortality in that village, but this positive impact of reducing under-five mortality comes entirely from the increase in the availability of improved toilets in the village. This claim is supported not only by the similarity between the magnitude of the positive impact of the percentage of households having toilets and the percentage of households having improved toilets but also by the plausible lack of impact of the percentage of households having non-improved toilets.

4.2.3. Robustness checks and a fallacy test

We conduct three robustness checks. First, in the main analyses, we dropped likely outlier villages where the total under-five mortality was above 15. However, this threshold may be too restrictive. Therefore, we modify this threshold to 20 and re-estimate the IV models. This includes 61 additional households in the sample. The results are qualitatively and quantitatively very similar to the main results.

Second, our main sample includes the villages in Phnom Penh, the capital city of Cambodia. Its population is about 2.3 million, which accounts for about 15% of the total population of Cambodia. The capital city differs substantially from other places in Cambodia in many aspects. Therefore, we exclude the sample of villages in Phnom Penh, and re-estimate the IV models. The results are qualitatively and quantitatively very similar to the main results, with greater statistical significance in the case of $Rnotoilet$, $Rtoilet$, and $Rimprovedtoilet$.

Third, we apply the inverse hyperbolic sine transformation to the slope variable and use it as the IV.[8] This modified IV is still strongly correlated with $Rnotoilet$, $Rtoilet$, and $Rimprovedtoilet$; the coefficient of the modified IV is statistically significant at the 1% level in the specification including the control variables, although the values of the F test of the excluded instrument are around 10. The second-stage results are qualitatively similar to the main results, although the coefficients of $Rnotoilet$, $Rtoilet$, and $Rimprovedtoilet$ for girls' under-five mortality become marginally insignificant (z-values are around 1.45).

Finally, as a fallacy test, we use mortality rate of children aged 6–10 years (total, boys and girls) as the outcome variables. However, we do not get any statistically significant coefficients of the toilet variables in the specification using the instrumental variable estimations.

Therefore, our main results hold robustly and improved toilets are considered to be effective in reducing mortality under the age of 5 years, but not above. These additional results facilitate the discussion in the next section. [9]

5. Discussion

According to our estimation results, the popularization of toilets in a given village reduces total, boys', and girls' under-five mortality in that village, and this positive impact of reducing under-five mortality comes entirely from the increased availability of improved toilets in the village. Conversely, the increase in the percentage of households without toilets increases total, boys', and girls' under-five mortality in the village. The family members of those households are thought to engage in open defecation. This practice contaminates drinking water sources, rivers, and food crops, leading to diseases, stunted growth, chronic malnutrition, and in some cases, death. Our estimation results clearly identify the worst-case scenario: under-five mortality. We have no clue whether the increased availability of non-improve toilets within a village affects the under-five mortality of the village or not.

Furthermore, it should be noted that in our analysis, positive and negative externalities regarding toilet issues within a village are fully taken into account because the sample unit is "village." Therefore, when a household newly constructs an improved toilet, it may positively impact the health condition of neighboring households and vice versa. Or, if a family member of a household without a toilet engages in open defecation, it may cause health problems for neighboring households and vice versa. Our analyses address these positive and negative externalities at the village level.

These findings have at least two policy implications. First, the policymakers must

---

[8] See Footnote 7 for inverse hyperbolic sine transformation.

[9] The estimation results for all the robustness checks and of the fallacy test are available upon request from the authors.

seriously consider the negative impacts of open defecation on under-five mortality in order to save young children's lives. A family member of a household without a toilet is highly likely to engage in open defecation. Accordingly, preventing open defecation must be a priority to save young children's lives.

Second, one of the recommended policy measures to save young children's lives and to prevent open defecation at the same time is to construct toilets in households. However, given the estimation results from our analyses, these need to be improved toilets. Furthermore, such policy measures would be more effective if implemented following a "big push" approach (Rosenstein-Rodan 1943). In other words, targeting the whole village, rather than individual households, would yield better outcomes. This is because such a large-scale intervention can leverage the positive externalities of constructing improved toilets. However, this requires a large amount of resources (i.e., money), and resource constraints are often a big challenge for policymakers in developing countries.

6. Conclusion

Using village-level balanced panel data of Cambodia, we investigate how the availability of toilets in a given village impacts under-five mortality in that village. The findings have at least two implications. First, an increase in the percentage of households without toilets increases total, boys', and girls' under-five mortality in the village. Second, only improved toilets appear to reduce under-five mortality, and whether non-improved toilets can reduce under-five mortality or not is inconclusive. We emphasize that the positive impacts of improved toilets are likely not the exclusive domain of under-five mortality, but might also include the reduction of diarrhea, stunted growth, and so on. Unfortunately, we were unable to investigate the impact of improved toilets on such outcomes due to a lack of data. However, such improvements are expected to yield better outcomes beyond under-five mortality.

Finally, we point out some limitations of this study. First, we utilized the data from two censuses in order to create the village-level panel data (2008 and 2019). The censuses were conducted 11 years apart, while the information on deaths covers only the 12 months prior to the census night, ignoring the deaths that occurred between 2009 and 2017. If we had had more detailed information for each year, we could have performed deeper analyses. However, in many developing countries, a census is conducted only every 10 years or so, and thus this limitation may be unavoidable.

Another important limitation is the external validity of the results obtained in the context of Cambodia. We believe the general pattern found in our study is interpretable, but a different setting with a different culture at a different stage of economic development may yield a different set of results. Therefore, similar analyses in different contexts are warranted.

References

Agha, Sohail. "The determinants of infant mortality in Pakistan." *Social Science & Medicine* 51.2 (2000): 199-208.

Anteneh, Andualem, and Abera Kumie. "Assessment of the impact of latrine utilization on diarrhoeal diseases in the rural community of Hulet Ejju Enessie Woreda, East Gojjam Zone, Amhara Region." *Ethiopian Journal of Health Development* 24.2

(2010).

Barnard, Sharmani, et al. "Impact of Indian Total Sanitation Campaign on latrine coverage and use: a cross-sectional study in Orissa three years following programme implementation." *PloS one* 8.8 (2013): e71438.

Cameron, Lisa, Susan Olivia, and Manisha Shah. "Scaling up sanitation: evidence from an RCT in Indonesia." *Journal of Development Economics* 138 (2019): 1-16.

Cameron, Lisa, et al. "Sanitation, financial incentives and health spillovers: a cluster randomised trial." *Journal of Health Economics* 77 (2021): 102456.

Chen, Baoqi, Fulei Jin, and Yaxin Zhu. "The impact of access to sanitary toilets on rural adult residents' health: Evidence from the China family panel survey." *Frontiers in Public Health* 10 (2022): 1026714.

Cheng, Shikun, et al. "Toilet revolution in China." Journal of environmental management 216 (2018): 347-356.

da Silva Wells, Carmen, and Christine Sijbesma. "Practical innovations for strengthening Community-Led Total Sanitation: selected experience from Asia." *Development in Practice* 22.3 (2012): 417-426.

Geruso, Michael, and Dean Spears. "Neighborhood sanitation and infant mortality." *American Economic Journal: Applied Economics* 10.2 (2018): 125-162.

Graham, Jay P., and Matthew L. Polizzotto. "Pit latrines and their impacts on groundwater quality: a systematic review." *Environmental health perspectives* 121.5 (2013): 521-530.

Kiviet, Jan F. "Testing the impossible: Identifying exclusion restrictions." *Journal of Econometrics* 218.2 (2020): 294-316.

Ministry of Planning of Cambodia 2020 "General Population Census of the Kingdom of Cambodia 2019: National Report on FINAL CENSUS RESULTS" available at https://www.nis.gov.kh/nis/Census2019/Final%20General%20Population%20Census%202019-English.pdf (last accessed on March 10, 2024)

Ndoziya, Alfonse Tapera, Zvikomborero Hoko, and Webster Gumindoga. "Assessment of the impact of pit latrines on groundwater contamination in Hopley Settlement, Harare, Zimbabwe." *Journal of Water, Sanitation and Hygiene for Development* 9.3 (2019): 464-476.

Rosenstein-Rodan, Paul N. "Problems of industrialisation of eastern and south-eastern Europe." *The economic journal* 53.210-211 (1943): 202-211.

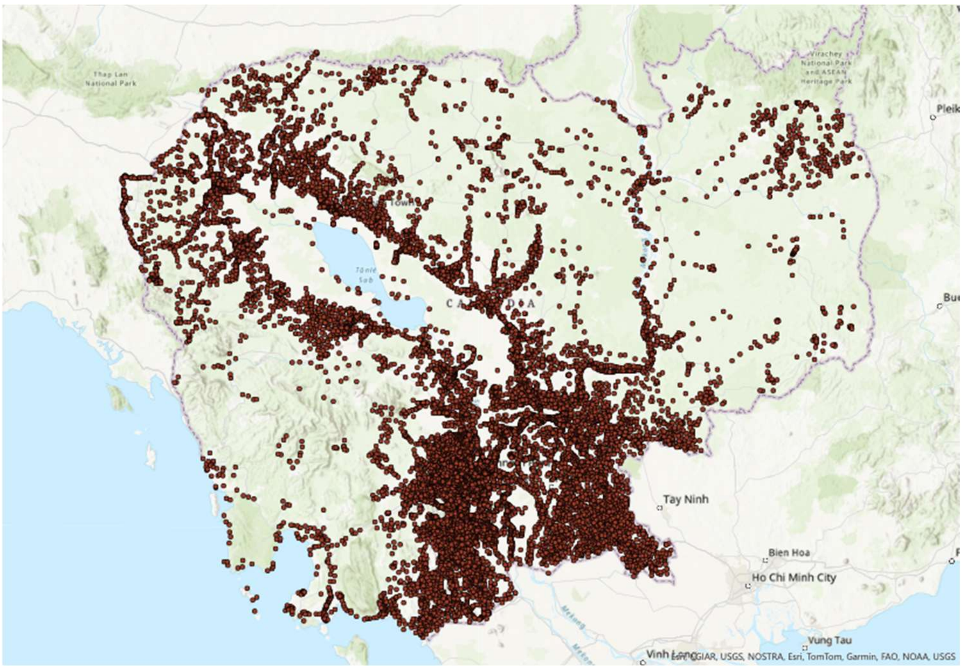
Semba, Richard D., et al. "Relationship of the presence of a household improved latrine with diarrhea and under-five child mortality in Indonesia." *The American journal of tropical medicine and hygiene* 84.3 (2011): 443.

Spears, Dean. "Effects of rural sanitation on infant mortality and human capital: Evidence from India's Total Sanitation Campaign." *Washington, DC: Princeton University* (2012).

Spears, Dean. "Exposure to open defecation can account for the Indian enigma of child height." *Journal of Development Economics* 146 (2020): 102277.

Wang, Dongqin, and Yanni Shen. "Sanitation and work time: Evidence from the toilet revolution in rural China." *World Development* 158 (2022): 105992.

Figure 1: Location distribution of villages in Cambodia



Each dot is the location of villages.
Source: The National Institute of Statistics of Cambodia

Table 1: Descriptive statistics

Variable	Explanation	Summary
Rdeathunder5	Total mortality of children under five years old (per 100 children)	0.656 (1.580)
Rdeathunder5M	Mortality of boys under five years old (per 100 children)	0.544 (1.855)
Rdeathunder5F	Mortality of girls under five years old (per 100 children)	0.793 (2.072)
Rnotoilet	The percentage of household having a toilet at home	46.411 (37.083)
Rtoilet	The percentage of household having no toilet at home	53.352 (37.124)
Rimprovedtoilet	The percentage of household having a non-improved toilet at home	46.350 (36.534)
Rnotimprovedtoilet	The percentage of household having an improved toilet at home	7.002 (15.344)
slope	The difference between the highest and lowest altitudes in the village (in meters)	12.843 (13.391)
RF_HHH	The percentage of households whose head is female	15.995 (10.423)
lNpop	Population of the village (after IHS transformation)	7.341 (0.650)
RWat_well	The percentage of households whose main water source is from a well	41.704 (39.600)
RWat_pond	The percentage of households whose main water source is from a pond	13.039 (26.574)
RWat_river	The percentage of households whose main water source is from a river	11.905 (27.107)
RM_ILT15_17	The percentage of illiterate males among aged 15–17 years	4.806 (12.488)
RF_ILT15_17	The percentage of illiterate females among aged 15–17 years	4.841 (12.724)
RM_ILT18_24	The percentage of illiterate males among aged 18–24 years	7.624 (16.263)
RF_ILT18_24	The percentage of illiterate females among aged 18–24 years	7.972 (16.945)
RM_ILT25_35	The percentage of illiterate males among age 25–35 years	10.715 (18.773)
RF_ILT25_35	The percentage of illiterate females among age 25–35 years	11.559 (19.854)
RM_ILT36_45	The percentage of illiterate males among age 36–45 years	14.526 (20.918)
RF_ILT36_45	The percentage of illiterate females among age 36–45 years	15.675 (22.226)
RZ_Fib_R_Elec	The percentage of households having a zinc or fibro roof house with electricity	20.195 (26.166)
RTil_R_Elec	The percentage of households having a tiled roof house with electricity	14.466 (21.960)
RNUM_FERT	The percentage of households using chemical fertilizer	58.286 (38.170)
RNUM_PEST	The percentage of households using pesticide	38.376 (38.093)
RFamily_Car	The number of family cars per household	0.031 (0.086)
lkmhe	Distance (km) from the village center to the health center (after IHS transformation)	1.775 (0.974)
Sample size		24,178

Mean (standard deviation)

IHS stands for “inverse hyperbolic sine.” See Footnote 7 in the text for more information.

Table 2: OLS estimation results on the relationship between toilet availability and under-five mortality

VARIABLES	(1) Total	(2) Boy	(3) Girl	(4) Total	(5) Boy	(6) Girl	(7) Total	(8) Boy	(9) Girl	(10) Total	(11) Boy	(12) Girl
Rnotoilet	0.0016*	0.0023***	0.0013	0.0017*	0.0032***	0.0006						
	(0.0008)	(0.0009)	(0.0011)	(0.0009)	(0.0010)	(0.0012)						
Rtoilet							-0.0015*	-0.0023***	-0.0012	-0.0017*	-0.0032***	-0.0005
							(0.0008)	(0.0009)	(0.0011)	(0.0009)	(0.0010)	(0.0012)
Village fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	No	No	Yes	Yes	Yes	No	No	No	Yes	Yes	Yes
Observations	24,178	24,178	24,178	24,178	24,178	24,178	24,178	24,178	24,178	24,178	24,178	24,178
R-squared	0.013	0.061	0.004	0.020	0.066	0.009	0.013	0.061	0.004	0.020	0.066	0.009

VARIABLES	(13) Total	(14) Boy	(15) Girl	(16) Total	(17) Boy	(18) Girl	(19) Total	(20) Boy	(21) Girl	(22) Total	(23) Boy	(24) Girl
Rimprovedtoilet	-0.0018**	-0.0014	-0.0025**	-0.0026***	-0.0028***	-0.0028**						
	(0.0008)	(0.0009)	(0.0011)	(0.0009)	(0.0010)	(0.0012)						
Rnotimprovedtoilet							0.0003	-0.0019	0.0022	0.0012	-0.0010	0.0032**
							(0.0011)	(0.0013)	(0.0014)	(0.0011)	(0.0013)	(0.0015)
Village fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	No	No	Yes	Yes	Yes	No	No	No	Yes	Yes	Yes
Observations	24,436	24,434	24,434	24,319	24,317	24,317	24,178	24,178	24,178	24,178	24,178	24,178
R-squared	0.013	0.061	0.004	0.020	0.066	0.010	0.013	0.061	0.004	0.020	0.065	0.009

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Total: total mortality rate of children under five years old
Boy: mortality of boys under five years old
Girl: mortality of girls under five years old

Table 3: IV estimation results on the relationship between toilet availability and under-five mortality

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
VARIABLES	Total	Boy	Girl	Total	Boy	Girl	Total	Boy	Girl	Total	Boy	Girl
Panel A: Second stage												
Rnotoilet	0.0135	0.0173	0.0102	0.0221**	0.0237**	0.0214*						
	(0.0115)	(0.0132)	(0.0139)	(0.0097)	(0.0111)	(0.0117)						
Rtoilet							-0.0132	-0.0169	-0.0099	-0.0218**	-0.0234**	-0.0211*
							(0.0112)	(0.0129)	(0.0136)	(0.0096)	(0.0110)	(0.0115)
Panel B: First stage												
		Rnotoilet			Rnotoilet			Rtoilet			Rtoilet	
slope		-0.2383***			-0.2927***			0.2442***			0.2962***	
		(0.0327)			(0.0303)			(0.0325)			(0.0302)	
F test of excluded instrument		53.00			93.17			56.41			96.05	
Village fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	No	No	Yes	Yes	Yes	No	No	No	Yes	Yes	Yes
Observations	24,178	24,178	24,178	24,178	24,178	24,178	24,178	24,178	24,178	24,178	24,178	24,178

Robust standard errors in parentheses

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

Total: total mortality rate of children under five years old

Boy: mortality of boys under five years old

Girl: mortality of girls under five years old

Table 4: IV estimation results on the relationship between toilet availability by type and under-five mortality

VARIABLES	(1) Total	(2) Boy	(3) Girl	(4) Total	(5) Boy	(6) Girl	(7) Total	(8) Boy	(9) Girl	(10) Total	(11) Boy	(12) Girl
Panel A: Second stage												
Rimprovedtoilet	-0.0119	-0.0152	-0.0090	-0.0227**	-0.0244**	-0.0220*						
	(0.0101)	(0.0116)	(0.0122)	(0.0100)	(0.0114)	(0.0120)						
Rnotimprovedtoilet							0.1217	0.1557	0.0918	-0.5475	-0.5873	-0.5303
							(0.1543)	(0.1890)	(0.1516)	(1.1858)	(1.2736)	(1.1627)
Panel B: First stage												
		Rimprovedtoilet			Rimprovedtoilet			Rnotimprovedtoilet			Rnotimprovedtoilet	
slope		0.2707***			0.2844***			-0.0264			0.0118	
		(0.0314)			(0.0286)			(0.0249)			(0.0251)	
F test of excluded instrument		74.34			98.40			1.13			0.22	
Village fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	No	No	Yes	Yes	Yes	No	No	No	Yes	Yes	Yes
Observations	24,178	24,178	24,178	24,178	24,178	24,178	24,178	24,178	24,178	24,178	24,178	24,178

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Total: total mortality rate of children under five years old
Boy: mortality of boys under five years old
Girl: mortality of girls under five years old