Institute for Fiscal Studies

Sanitation and child health in India

IFS Working Paper W15/32

Britta Augsburg
Paul Rodriguez-Lesmes

1965 - 2015 50 E·S·R·C 50 YEARS SHAPING SOCIETY

The Institute for Fiscal Studies (IFS) is an independent research institute whose remit is to carry out rigorous economic research into public policy and to disseminate the findings of this research. IFS receives generous support from the Economic and Social Research Council, in particular via the ESRC Centre for the Microeconomic Analysis of Public Policy (CPP). The content of our working papers is the work of their authors and does not necessarily represent the views of IFS research staff or affiliates.

Sanitation and child health in India

Britta Augsburg (IFS) and Paul Rodríguez-Lesmes (UCL)*

1st December 2015

Abstract

Our study contributes to the understanding of key drivers of stunted growth, a factor widely recognized as major impediment to human capital development. Specifically, we examine the effects of sanitation coverage and usage on child height for age in a semi-urban setting in Northern India. We use instrumental variables to control for endogeneity of sanitation usage coverage. We find that sanitation coverage plays a significant and positive role in height growth during the first years of life.

JEL I12

Keywords: child health; sanitation coverage; open defecation; India.

*We thank the ESRC-DfID (grant number ES/J009253/1) as well as FINISH Society for financial assistance and Sanghmitra Gautam for allowing access to the price data. All errors that remain our own. Author affiliations and contacts: Augsburg (Institute for Fiscal Studies, britta_a@ifs.org.uk); Rodríguez-Lesmes (University College London, p.lesmes.11@ucl.ac.uk).

1 Introduction

The failure to reach linear growth potential early in life has been widely recognized as a major impediment to human capital development. There is increasing evidence that growth failure (stunting), often associated with poor nutrition[1], is correlated, likely in a causal way, with lower educational and labour market attainments (Victora et al. [2010], Behrman et al. [2009], Hoddinott et al. [2008, 2013], Maluccio et al. [2009]). Rates of stunting, the general term for a child being short for its age, have been reducing over recent years, but 159 million children around the world are still estimated to be affected, more than half of these living in Asia (de Onis et al. [2015]).

While a growing body of literature is contributing to our understanding of the consequences of stunting, knowledge is still limited with respect to the key drivers of low height for age. It is generally understood that inadequate diet and diseases are important immediate causes of stunting (Black et al. [2008], Smith and Haddad [2014], Bozzoli et al. [2009]) but dealing with the endogeneity of these inputs remains a challenge in the literature (Deaton [2007]).

In this study we focus on the role of diseases, specifically the role of an improved disease environment in the growth trajectory of children under the age of 5 years. Diseases have been linked to stunting (Checkley et al. [2008]) but have also shown direct associations with short (Nokes et al. [1992], Nokes et al. [1998], Walker et al. [2011]) and long-term effects on human capital (Almond and Currie [2011], Bozzoli et al. [2009]). Understanding the potential of improving the disease environment that children live in is hence of direct policy relevance.

The disease that is identified to be of primary concern is diarrhoea. WHO acknowledges diarrhoea to be the leading cause of child mortality and morbidity in the world, killing an estimated 760,000 children every year [WHO, 2013]. Most of these diarrhoea cases are believed to be due to contamination of the environment. Eighty percent are seen to be linked to unsafe water, inadequate sanitation or insufficient hygiene as estimated in a 2008 report by the WHO [Pruesss-Uestuen et al., 2008]. Effective and affordable interventions that aim to improve the disease environment by tackling access to safe drinking water, adequate sanitation and hygiene behaviour have therefore been an important focus on the developing agenda.

Rigorous evidence on the potential of improvements in water access and quality exists and the complementarities between water and sanitation have been recognized [WHO,

[1] The WHO describes stunted growth (low height-for-age) as "a process of failure to reach linear growth potential as a result of suboptimal health and/or nutritional conditions".

2008, Duflo et al., 2015b, Pruesss-Uestuen et al., 2008]. The understanding of the role of sanitation itself is however still limited. Recent randomised controlled trials fail to show causal impacts of improvements in sanitation and child health [see for example Clasen et al. [2014][2]. Hypothesised reasons for this are manifold and mostly link to technological, financial and behavioural challenges. Such limited understanding and evidence of effectiveness is particularly problematic for an investment that faces significant challenges, including lack of appropriate technology, local capacity and most importantly, lack of financial resources - and with that is easily discouraged.

We provide evidence on causal links between an improved sanitation environment and child health, proxied by height for age of children using two rounds of primary data collected for learning around a sanitation intervention in Northern India. We follow recent advances in the sanitation literature, concentrating on sanitation coverage and usage as our main variable of interest. A number of studies are currently trying to link rates of toilet ownership or open defecation to diarrhoea or child height for age. Many of these are in initial stages and hence not published at the time of writing, potentially driven by their limited ability to account for endogeneity of this variable of interest.

The main motivation lies in the understanding that individual household sanitation is unlikely to live up to promises in improving health statuses when neighbours are still contaminating the environment, i.e. that externalities are at play. The percentage of households in a community rather than private ownership is hence hypothesised to be the more relevant unit of observation when analysing health impacts.

Using an instrumental variable approach to account for endogeneity of community latrine use, our approach is closest to the currently unpublished work by Geruso and Spears [2014], which uses the fraction of Muslims in a village as an instrument. Acknowledging themselves the 'important possibility of bias' since 'Muslim concentration is not perfectly uncorrelated with observable characteristics that plausibly impact child health', they find that a decline in 2.6-2.9 infant deaths out of 1,000 can be achieved with a 10% reduction in the fraction of neighbours defecating in the open. Gertler et al. [2014] also use an instrument in estimating the impact of open defecation rates on child health (measured by child height). They exploit random allocation of sanitation intervention in their data set. However, it is unlikely that the interventions impacted child health only through reducing open defecation rates, given that intervention activities included for example hygiene behaviour campaigns. Their suggested impact of a one standard deviation reduction in their constructed open defecation index would lead to an average increase of standard deviation in children's height. Finally, Andres et al.

[2] An exception is a recent study published in the Lancet which analyses a community-led sanitation intervention in remote areas of Mali [Pickering et al., 2015].

[2014], use a simple cross-sectional approach, not attempting to account for endogeneity in their variables of interest, finding that 'a 47 percent reduction in diarrhoea prevalence between children living in a household without access to improved sanitation in a village without coverage of improved sanitation and children living in a household with access to improved sanitation in a village with complete coverage'.

Our identification strategy improves on these papers by following the production function literature in identifying unbiased estimates. In this literature, prices are typically acknowledged to affect investment choices without entering the production function directly (Todd and Wolpin [2003], Puentes et al. [2014], Attanasio et al. [2015]). Our data includes prices on raw materials of toilets, which exhibit sufficient geographic variation and explanatory power to serve as a useful instrument. Applying this estimation approach to our data, suggests that a ten percent increase in households using private or community toilets in (semi-) urban slum communities of India increases child height for age z-scores significantly by 0.15 standard deviations. Sub-sample analysis suggests that this effect is primarily driven by female children.

We will describe the estimation approach, as well as the limitations we face, in the next section. We then introduce the data and its context in Section 3, before presenting and discussing our findings in Section 4. Section 5 concludes.

2 Methodology

2.1 A simple economic model for determinants of child health

The principal objective is to understand the relationship between community level sanitation and child health, acknowledging that behavioural responses can induce endogeneity in our variable of interest.

Mosley and Chen [1984] suggested a useful framework for us to built on. Their framework, which integrates approaches from demographers and epidemiologists, identifies a set of exogenous and endogenous determinants of child health and survival, acknowledging the role of household and community sanitation in determining child health and survival. Factors identified as exogenous include individual and household characteristics such as maternal education, income and family composition, institutional factors such as community infrastructure, ecological factors, such as rainfall and cultural factors, such as traditions and norms. Factors identified as endogenous are referred to as proximate determinants and include breastfeeding and household sanitation ownership. Combining insights from this framework with those from recent advances in the understanding of human capital production functions is useful in guiding the choice of variables to include in the estimation, in understanding which variables are likely to be endogenous and whether important determinants have been omitted. We extend Currie [2000] economic model of the determinants of child health in such direction. In this unitary household model, parents maximize the following objective function:

$$\sum_{t=1}^{T} E_t \beta^t U_{ivt} + B(A_{iv,T+1})$$

In this one-child household i in village v, inhabited by I_v households, parents are altruistic and get utility from their children's health status and the bequest, B, they leave to them. Period-specific utility is given by:

$$U_{ivt} = U(Q_{ivt}, S_{ivt}, C_{ivt}, L_{ivt}; X_{ivt}, u_{1iv}, \varepsilon_{1ivt})$$

where, Q_{ivt} is child health, C_{ivt} is other consumption, L_{ivt} is leisure; and taste for them might differ according to some observed (X_{ivt}) and unobserved characteristics (u_{1iv}) and shocks (ε_{1ivt}). On top of this, households get utility from having access to a sanitation facility S_{ivt}. Reasons for this direct benefit of sanitation might include comfort, social-status, security, as well as health considerations of the adults.

In the original model by Currie [2000], the evolution of child health is shaped by parental physical, G_{ivt}, and time investments, V_{ivt}. Their productivity depends on observed (Z_{ivt}) and unobserved (u_{2iv}) characteristics as well as unobserved shocks (ε_{2ivt}). We extend the model to include an additional element: namely, what we term, 'environmental sanitation'. $ES_{v,t} = ES_{v,t} = \frac{1}{N_v}\sum_{i=1}^{I_v} S_{i,v,t}$. We define this term, our main variable of interest (which we refer to as 'environmental sanitation') as the percentage of one-child households, i, in the village (including i) that use sanitation infrastructure for defecation, where $S_{i,v,t}$ is an indicator variable = 1 if all members of a randomly selected household in village v use the toilet they own or use a community toilet. The variable is zero otherwise. I_v indicates the total number of randomly selected households in village v.

This definition is driven by our interest in the role of infrastructure that isolates human waste, faeces, from the environment, i.e. sewage, community toilets and private household toilets. More specifically we are interested in the *usage* of such facilities[3], primarily private household sanitation, but also the less common usage of community toilets and neighbours' toilets. The Mosley and Chen [1984] framework defines sanitation ownership as a proximate determinant of child health, acknowledging its importance in providing a hygienic environment as well as the fact that it is likely to be endogenous. For example, endogeneity might stem from households with a child that has a particularly weak immune system possibly being more likely to seek investment in infrastructure that keep the household's imminent environment free from contaminants,

[3]One of the reasons put forward for non-impacts on health in for example the study by Clasen et al. [2014] is that the constructed toilets were not used.

contributing to a negative correlation between demand for curative health inputs and good health. This is in contrast to the anticipated positive relationship of improvements in the imminent disease environment and health if such an improvement were randomly allocated to households of equally weak children.

In our definition of disease environment we go beyond the imminent disease environment of the household, acknowledging that toilet ownership and usage provides a direct benefit as well as an external benefit, which is believed to be substantial (Duflo et al. [2015a]; Gertler et al. [2014], Geruso and Spears [2014], Andres et al. [2014]). Using a toilet reduces own contact with faeces in addition to other private benefits a toilet might provide (time saving, privacy, etc). It further reduces the rate of open defecation, what is believed to be a major cause for parasite infections and diarrhoea, particularly observed in children under five years of age.

It is hence not just one's own toilet usage behaviour that determines health, but also the behaviour of neighbours and community members. As is the case for the individual ownership of sanitation infrastructure and discussed by Mosley-Chen, also this broader definition of sanitation environment is likely to be endogenous. Take for example communities with very high population density and at the same time limited public (health) infrastructure, as often the case in for example slums in developing countries. One can imagine that communities faced with such conditions which are likely to negatively impact health, to be more likely to make their own investments in infrastructure improving the disease environment. We will therefore need to deal with the likely endogeneity of $ES_{v,t}$.

Including this variable in Currie [2000]'s model of child health, we get a health production function which is a function of sanitation coverage. In other words, one of the relevant determinants of child health is determined at the village level. Depending on $f(\cdot)$, individuals might control or not this input. As a result, the health production function takes the following structure:

$$Q_{ivt} = f(Q_{iv,t-1}, G_{ivt}, ES_{vt}; Z_{ivt}, u_{2iv}, \varepsilon_{2ivt}) \tag{1}$$

The rest of the model follows Currie's structure. Parents get income from working H_{ivt} hours (where available time is normalised to unity), which reduces the amount of time available for leisure as well as investments in the child's health. Physical resources are distributed among savings, child-investments, a one-off sanitation investment T_{ivt}, and consumption. Relative to the standarised prices of other consumption, prices of child investments P^G_{vt} and toilet construction P^T_{vt} determine the marginal cost of both investments. Notice that once a household builds a toilet, its sanitation environment is assumed to improve permanently in the following period, through the personal ownership as well as the externality effect. This reflects the fact that gains from sanitation might not be immediate. Such resources grow with income Y which can come either from work at a wage w, from capital rent at a rate r, or from other source I_{ivt}. The related equations are:

$$C_{ivt} = Y_{ivt} - P_{vt}^{G} G_{ivt} - P_{vt}^{T} T_{ivt} - (A_{t+1} - A_t)$$

$$Y_{ivt} = I_{ivt} + w_{vt} H_{ivt} + r A_t$$

$$L_{ivt} + V_{ivt} + H_{ivt} = 1$$

$$S_{ivt} = \max(S_{iv,t-1}, T_{iv,t-1})$$

The model can be solved, and as in the original setup, to yield Frish demand functions. Within these, λ denotes the marginal utility of wealth and M corresponds to a vector of moments of the distribution of future observed and unobserved variables $\{X_{iv\tau}, Z_{iv\tau}, P_{iv\tau}, \varepsilon_{2iv\tau}, \varepsilon_{1iv\tau}, S_{iv\tau}^{-i}\}_{\tau=t+1}^{T}$. Here, S_{ivt}^{-i} is a vector which incorporates the sanitation status of all other households in village v, and P_{vt} is a vector of prices (including wage) at the village level for a given period, t. [4]

The Frish demand functions are of the following form:

$$C_{ivt}, H_{ivt}, T_{ivt}, G_{ivt} \text{ and}$$
$$V_{ivt} = F(\beta, r, \lambda_{ivt}, X_{ivt}, Z_{ivt}, S_{ivt}, P_{vt}, ES_{vt}, u_{1iv}, u_{2iv}, \varepsilon_{2ivt}, \varepsilon_{1ivt}, M_{ivt}).$$

Given these, we can substitute both physical and time inputs into the health production function, Equation 1. If we also substitute for λ_{ivt} using the budget constraint, and assuming that M_{ivt} and A_{ivt} are functions of realizations of current, and past exogenous variables $J_{ivt} = \{X_{iv\tau}, Z_{iv\tau}, P_{v\tau}, I_{iv\tau}, \varepsilon_{2iv\tau}, \varepsilon_{1iv\tau}, S_{iv\tau}^{-i}\}_{\tau=1}^{t-1}$ and A_{iv0}, we get:

$$Q_{ivt} = f^{'}(Q_{iv,t-1}, A_{iv0}, \beta, r, I_{ivt}, , X_{ivt}, Z_{ivt}, S_{ivt}, P_{vt}, ES_{vt}, u_{1iv}, u_{2iv}, \varepsilon_{2ivt}, \varepsilon_{1ivt}, J_{ivt})$$

and iterating over Q results in Equation 2. This reduce form equation of the production function makes it clear that there is a link between sanitation prices and health. Such link arises due to the reduction on the marginal cost of building a toilet, which increases demand for such good.

$$Q_{ivt} = f^{'}(Q_{iv0}, A_{iv0}, \beta, r, I_{ivt}, , X_{ivt}, Z_{ivt}, S_{ivt}, P_{vt}, ES_{vt}, u_{1iv}, u_{2iv}, \varepsilon_{2ivt}, \varepsilon_{1ivt}, J_{ivt}) \tag{2}$$

A reduced form expression for toilet ownership can also be derived:

[4]Notice that if household i has an important weight in determining ES_{vt}, $S_{iv\tau}^{-i}$might be a function of $\{X_{iv\tau}, Z_{iv\tau}, P_{iv\tau}, \varepsilon_{2iv\tau}, \varepsilon_{1iv\tau}\}_{\tau=t+1}^{T}$. Given this, household's i best response implies that the demands should include moments for all future variables of all individuals in the village$\{\{X_{\iota v\tau}, Z_{\iota v\tau}, P_{\iota v\tau}, \varepsilon_{2\iota v\tau}, \varepsilon_{1\iota v\tau}\}_{\tau=t+1}^{T}\}_{\iota=1}^{I_v}$. Here, for simplicity, we assume that this household has virtually no power in determining everyone else's adoption decision and that $S_{iv,t+1}^{-i}$ can be forecasted with some village characteristics.

$$T_{ivt} = T(A_{iv0}, \beta, r, I_{ivt}, , X_{ivt}, Z_{ivt}, S_{ivt}, P_{vt}, ES_{vt}, u_{1iv}, u_{2iv}, \varepsilon_{2ivt}, \varepsilon_{1ivt}, J_{ivt})$$

As a result of the above, environmental sanitation at the village level is determined by a full set of present and past states $\theta_{vt} = \{Q_{iv0}, A_{iv0}, \{X_{iv\tau}, Z_{iv\tau}, I_{iv\tau}, \varepsilon_{2iv\tau}, \varepsilon_{1iv\tau}, S^{-i}_{iv\tau}\}^{t-1}_{\tau=1}\}^{I_v}_{i=1}$, which includes village level characteristics, and, importantly for our sub-sequent analysis, the village-specific vector prices.

$$ES_{iv,t+1} = f^{*}(\beta, r, S_{1vt}...S_{I_v vt}, \theta_{vt}, P_{vt})$$

The model shows us that both the health production function as well as the demand for toilet ownership are influenced by unobserved idiosyncratic persistent and transitory shocks, initial conditions, and by the history of exogenous variables which might only be partially unobserved. Our goal is to identify $E[\partial Q_{ivt}/\partial ES_{ivt}]$, and given the presence of confounders, we will identify such marginal effect by exploit village level variation of P_{vt}, which induces exogenous variation on ES_{vt}. Notice that an additional channel is still open: the functional form of $U(\cdot)$ might imply that the demand for physical investments might be directly affected by the price of sanitation, for instance, with a CES specification. Such effects are expected to operate in an opposite direction to ES, as lower prices of raw materials will induce less physical investments, reducing Q. If that is the case, our estimates would be provide a bound of the impact of environmental sanitation. Another issue is if ES and the other inputs are substitutes or complements in the production function, which will imply different allocation of the inputs given the exogenous variation on ES. In the most extreme scenario, all the impact on health would be driven by agents that invest more on their children under the believe that the productivity of such investments is going to increase. Such questions on the functional form are beyond the scope of this paper.

In order to provide an estimate of such impacts, given the limitations of the data, we will impose some restrictions. First, $P_{vt} = P_{vt-1}$, as we do not have variation in time of such vector. Second, we will assume that the relationship between environmental sanitation and prices is as good as linear, as well as between child health and environmental sanitation. These strong assumptions restrict the analysis and avoid potentially key elements as non-linearity between ES and Q. Nevertheless, they allow us to get an idea of the strength of the link between both variables.

2.2 Estimation specification

Taking this model to the data, our regression specification becomes:

$$Q_{i,v,t} = \alpha + \gamma ES_{v,t} + \delta_1 X^{c}_{i,v,t} + \delta_2 X^{hh}_{i,v,t} + \delta_3 X^{v}_{i,v,t} + \varepsilon^{Q}_{i,v,t} \tag{3}$$

with $X^{c}_{i,v,t}$ representing relevant individual, i.e. child-level, characteristics, include age, gender and whether the child was (or is) breastfed. Household level variables, $X^{hh}_{i,v,t}$, include the household composition, the education of the main woman in the household[5], income and shocks experienced; village level characteristics, $X^{v}_{i,v,t}$, such as information on water and garbage disposal; $\varepsilon_{i,v,t}$ are shocks to health in period t.

To address the endogeneity of our main variable of interest, $ES_{v,t}$, we employ an instrumental variable approach, estimating the following first-stage regression:

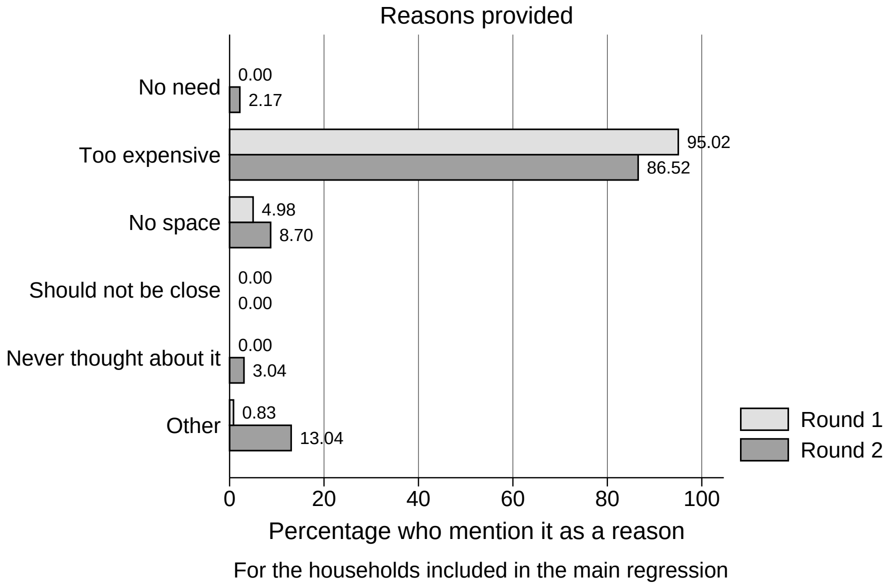
$$ES_{v,t} = \mu_0 + \mu_1 X^{c}_{i,v,t} + \mu_2 X^{hh}_{i,v,t} + \mu_3 X^{v}_{i,v,t} + \mu_2 Z_{v,t} + \varepsilon^{ES}_{i,v,t} \tag{4}$$

Our choices of instrument, $Z_{v,t}$, follows the production function literature. In this literature, prices are typically acknowledged to affect investment choices as also outlined in the model above, without entering the production function directly (Todd and Wolpin [2003], Puentes et al. [2014], Attanasio et al. [2015]), reflected in equation 2 above. Prices for sanitation raw materials are the key candidate. A common constraint in the use of these prices is limited variability. We benefit from prices that exhibit sufficient geographic variation at the community level to serve as instruments, while at the same time being a significant predictor of sanitation uptake as we will discuss in more detail below in Section 3.1.

In this approach, our parameter of interest, γ, is hence to be interpreted as a local average treatment effect. We are implicitly comparing the average level of child health for communities where dwellers are willing to build toilets but are restricted due to prices to those where the restriction does not apply. Considering the large percentage of households in our study sample that report financial constraints to be the main one to sanitation uptake, we believe that this is a reasonable approach to follow. Figure 1 shows the reported reasons why households do not own a toilet for our study population, which we will describe in more detail in the next section.

[5] One could argue that a better indicator to include would be the education level of the child's mother specifically. Our data does not allow to identify this relationship within the household. Given the household composition, we can infer that in many cases the main woman is likely to be the mother. Where it is not, it is likely that the practices by the mother are influenced by the main woman in the household.

Figure 1: Why not a toilet?



2.3 Limitations

We face a number of limitations driven primarily by our sample size and variables available in the data set.

The main constraint we face is that, although we have two rounds of data available, we are not able to use the data in a panel context. The reason for this is twofold: For one, the data collection was not designed to track individual households. Based on names available, we are able to imperfectly match individuals across rounds. However, given numerous ways of spelling names, not all children can be matched. More importantly, given that three years passed between the two survey rounds, many of our children at baseline were older than 5 years at follow-up and their height was hence not measured anymore in this second survey round. Due to this limitation, we are not able to include any child fixed effects in our specification.

Including such a fixed effect is often done to account for genetic endowment (see for example Puentes et al. [2014]). We are able to proxy for health endowment of the child by controlling for it's mother's height instead. This would primarily proxy for heritable endowment, which is seen as an important, unobserved determinant of child health. Medical papers suggest that 60-80 percent of height variation is determined by genetic factors (Ginsburg et al. [1998]; Silventoinen [2003]). Ideally, we would like to also include the height of the child's father. Unfortunately, anthropometrics of male adult household members were not collected. It is however quite common to use only one parent's measure as a proxy for inherited endowment[6].

[6]This is also the case in the literature on early childhood development and education production

Similarly, the fact that we are not able to link children across rounds restrains us from accounting for lagged height of the child, which would serve as a proxy for the full history of past inputs. It is not clear though whether its inclusion would be indeed beneficial given its likely endogeneity. The reason for this lies in parents likely response to health shocks. On the other hand, if our instrument is unrelated to this variable, our estimate is not biased due to its exclusion.

Another factor that is known to be an important determinant of child health is of course nutrition. We have limited information in our data on food items consumed by children in our study households in particular. We are able to construct a diversity index for children older than two years of age, including this however comes with a number of sacrifices. For one, excluding children age 0-2 years from the analysis means excluding those children where the plasticity in growth is seen most important (Victoria et al. [2010]). Further, not only would our sample size, and with that our power, be considerably reduced, but we are also faced with important non-reporting which seems to be systematically related to household and child characteristics (age and gender of the child, age of the mother, and household size). We are therefore not able to show any meaningful results which include information on child nutrition in our analysis.

3 Data and context

The context of our study are households residing in slums and peripheral villages of the city Gwalior. Gwalior is a historical and major city in the state of Madhya Pradesh, India, with an estimated slum population of one fourth of its citizens (Aggarwal and Kumar [2008]). This puts Gwalior above the country average of about 17% of urban households living in slums according to the 2011 slum census. This is an important population to study since on the one hand, they typically live in very crowded conditions, implying more important sanitation externality links (see for example Hathi et al. [2014]), while at the same time experiencing on average worse access to sanitation than the already low national average. The 2008-09 National Sample Survey Organisation [NSSO, 2010] survey estimates that 81 per cent of slum-dwellers in India have inadequate access to sanitation, which compares to national urban sanitation coverage rates of 26% in 2011.

At the same time, Madhya Pradesh is amongst states experiencing the worst rates of underweight and stunting for children. A nationwide survey, the Rapid Survey on Children (RSoC), conducted in 2013-14 by the Ministry of Women and Child Development in cooperation with UNICEF, revealed that a staggering 44.7 percent of children under 5 years of age were stunted (18.5% severely stunted) in Madhya Pradesh compared to a national average of 38.7%.

functions, where for example the mother's mother's AFQT score is commonly used to proxy for genetic endowment of the child (Todd and Wolpin [2003]).

The data we use in this study was collected with the intention of evaluating a sanitation called FINISH intervention implemented by the voluntary organisation Sambhav in Gwalior (Madhya Pradesh).[7] The intervention focused primarily on the uptake of private household sanitation.[8] The evaluation design allocated 39 slums and 17 peripheral villages (henceforth we will refer to them jointly as communities) to be exposed to the sanitation intervention and a control group. However, due to challenges faced by the implementing partner, primarily due to the Indian microfinance crisis, the intervention was on the one hand implemented at a much lower intensity than initially envisioned and on the other hand, the treatment allocation was not adhered to, so that we are not able to benefit from such endogenous variation. Baseline (BL) and follow-up (FU) surveys were nevertheless implemented. The BL survey was conducted between February and April 2010, and the FU survey between March and December 2013. In total, 1,982 households (HHs) were interviewed at BL, covering 11,032 individuals. These households were a representative sample of the community at that time. For the FU survey 2,020 HHs were interviewed, covering 12,360 individuals. 1,816 of these 2,020 HHs are in both BL and FU, the remaining were included as a replacement sample. The attrition of panel households is hence 8%.

Our observations in this analysis are children that are 5 years or younger at baseline (359 children) and follow-up (605 children), providing us with a sample of 964 children. As discussed above, we are not able to link the individuals across surveys and are hence not able to look at panel specifications.

Table 1 provides information on key characteristics of our sample children. The information under 'BASIC' are characteristics for all children in the communities we analyse in this study. Statistics under 'MAIN' refer to our sample from the regression analysis. We lose children due to some missing characteristics. These can be different variables for different children, which do not appear to be due to any systematic response rates.

As Table 1 further reveals, slightly less than half of the sample children are female (∼48%) and the average child in our sample is 35 months old, and lies below the reference population for both weight for age and height for age. Average weight for age z-scores, our outcome variable, is -1.7 on average in our population, indicating that the average child is stunted with respect to the reference population. If our children were on track with respect to growth given their age, the average expected value would be zero.

[7]Sambhav engages in issues of women empowerment, health, sanitation, education and violence against women, through direct program intervention and policy level advocacy. They work in about 1,500 villages/slums in Madhya Pradesh and Uttar Pradesh, primarily with Sahariya Adivasis, women and dalits, children and the disabled.

[8]Details on the intervention can be found in the endline report available from the authors upon request.

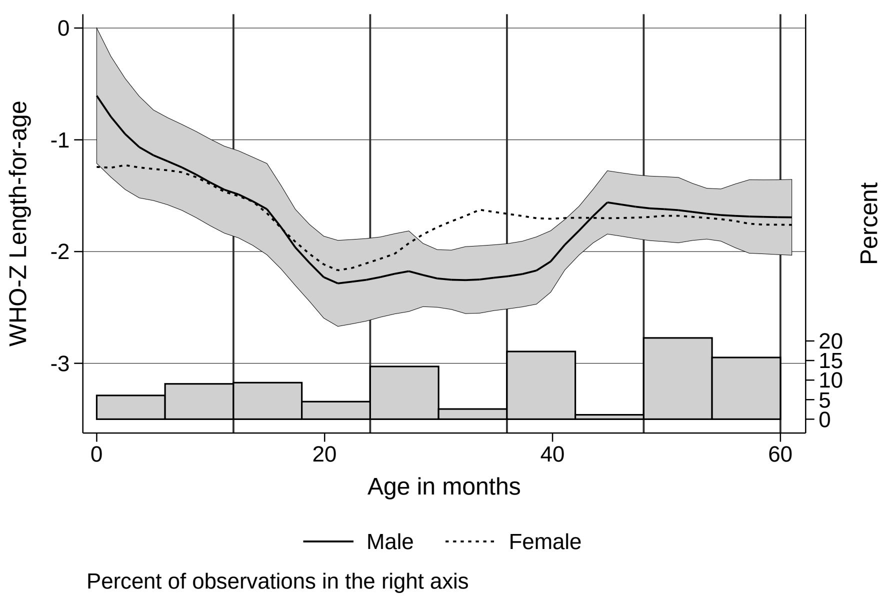
Table 1: Descriptive Statistics Children

	BASIC				MAIN			
	Mean	SD	Min	Max	Mean	SD	Min	Max
Children								
Age in months	34.5	18.2	0.5	61.0	34.8	18.2	0.5	61.0
Female	48.2%				48.4%			
Weight-for-age z-score	-1.7	1.6	-6.0	5.5	-1.7	1.6	-6.0	5.5
Length/height-for-age z-score	-1.6	2.2	-6.0	6.0	-1.6	2.2	-6.0	6.0
Weight-for-length/height z-score	-1.1	1.7	-5.8	5.9	-1.1	1.7	-5.8	5.9
BMI-for-age z-score	-0.9	1.8	-6.0	5.5	-0.9	1.8	-6.0	5.5
Total Children Round 1	358				331			
Total Children Round 2	605				532			
Total Children	963				863			

BASIC: For which there is information about children and main woman age, height, gender, and that live in a village where price data was collected. **MAIN:** Same sample as the main regressions

We picture the WHO height for age z-score for our sample children in Figure 2. It shows the z-score by age of the child, with the bars at the bottom providing information on the age distribution in our sample. We can see that the z-score reduces particularly in the first two years of life, after which children seem to catch up slightly again, staying still far from the standard population though. This trend is similar for boys (solid line) and for girls (dotted line).

Figure 2: Z-LEN by Gender



The primary survey data further includes detailed information on the households socio-economic status which will be important to account for in our analysis as discussed in the section on methodology. We for example include information on the households' size and composition as well as wealth levels (as measured by income) and education levels of the main woman in the household. These are included under the presumption that they are correlated with omitted inputs and will alleviate omitted variables bias. Table 2 provides descriptive statistics of household and community level characteristics.[9]

About one fourth of sample households are Muslim, remaining are Hindu and only for a very small percentage is the religion unknown[10]. In terms of the caste, 20-30% of sample households are scheduled tribes or schedule castes and another 5% are other minority backward castes. Almost 20% of households report in the interview that they belong to the forward caste.

Our households of interest consist of on average 6-7 household members of which half are male. The average annual income is around 70,000 Rupees, which is about USD 2,000, implying that these households live on average on USD 3 per day. Putting this number in the context of the household size, it is clear that these households are much below the internationally used poverty line of USD 1.25 per person per day. Not surprisingly then do we find that the total consumption expenditures exceed the household income. Partly this is driven by the fact that consumption expenditures include value of home produced and traded food. Most households do however live in a dwelling of strong or semi-strong structure. This, at least partly, reflects that the slums included in the study are registered slums with primarily strong dwelling structures. On average in the two survey rounds, almost 50% of these households owned a toilet.

As for the main woman of the household, we can see that more than half do not have any education and is on average 30 years of age. They are on average 1,50m tall.

[9]Monetary values are in Indian Rupees of 2013: R1 values where adjusted by a factor of 1.32. It was calculated based on national level gures for 2011, 2012 and 2013.(†)

[10]We include this 'unknown' variable in our analysis so not to loose these observations while at the same time being able to account for the religion which has been shown an important determinant of sanitation behaviour (Geruso and Spears [2014]).

Table 2: Descriptive Statistics Households

	BASIC				MAIN			
	Mean	SD	Min	Max	Mean	SD	Min	Max
Social background								
=1 if Muslim	22.4%				24.1%			
=1 if unkown religion	0.7%				0.5%			
=1 if forward caste	18.6%				18.7%			
=1 if minority backward caste	5.2%				5.7%			
=1 if scheduled caste or tribe	28.5%				28.7%			
=1 if unkown Caste	7.4%				6.7%			
HH Characteristics †								
Number of HH members	6.6	2.5	3.0	15.0	6.5	2.4	3.0	15.0
Number of male HH members	3.3	1.6	1.0	11.0	3.2	1.6	1.0	11.0
Any household shock last 12 months	9.8%				8.3%			
HH self-reported Income	70.1	47.6	0.0	276.0	69.0	47.1	0.0	276.0
Total consumption expenditures of hh in last year	101.8	81.9	6.1	704.0	97.6	76.3	6.1	704.0
Income: 0.00 - 40.00 K Rup	24.9%				24.9%			
Income: 40.15 - 60.00 K Rup	31.6%				32.7%			
Income: 60.25 - 90.00 K Rup	19.8%				19.8%			
Income: 90.30 - 280.00 K Rup	23.7%				22.5%			
Type of dwelling: strong	59.2%				56.7%			
Type of dwelling: semi-strong	30.7%				33.0%			
Type of dwelling: other	10.1%				10.3%			
Mother characteristics								
Education: no formal	56.4%				56.8%			
Education: 1-5 yrs	14.2%				14.0%			
Education: 6-8 yrs	16.1%				16.7%			
Education: 9 yrs +	13.2%				12.5%			
Age (Yrs)	31.6	10.1	7.0	85.0	31.4	10.1	7.0	85.0
Height (cm)	149.6	6.7	114.0	194.0	149.6	6.6	116.0	194.0
Sanitation and Hygiene								
Village/slum is a slum	59.6%				62.2%			
Baseline Observation	37.7%				39.2%			
=1 if HH has a toilet	48.7%				48.0%			
Uses a toilet	47.1%				46.9%			
Total Households	298				277			
Households Round 1	266				247			
Households Round 2	440				383			

BASIC: For which there is information about children and main woman age, height, gender, and that live in a village where price data was collected. **MAIN:** Same sample as the main regressions

Table 3: Descriptive Statistics Communities

	BASIC				MAIN			
	Mean	SD	Min	Max	Mean	SD	Min	Max
Village Level								
Village/slum is a slum	69.8%				71.1%			
Cost of Raw Materials (1000s Rps)	8.3	1.7	5.5	10.8	8.4	1.6	5.5	10.8
Round 1: Village: % Toilet Ownership)	41.7	32.8	0.0	100.0	41.8	33.3	0.0	100.0
Round 1: Village % who uses a toilet	44.5	35.3	0.0	100.0	44.7	35.7	0.0	100.0
Round 1: Village has sewage	47.2%				48.6%			
Round 2: Village: % Toilet Ownership)	59.0	27.7	0.0	100.0	58.1	29.1	0.0	100.0
Round 2: Village % who uses a toilet	56.6	30.4	0.0	100.0	55.6	31.8	0.0	100.0
Round 2: Village has sewage	77.1%				76.7%			
Villages Round 1	38				37			
Villages Round 2	43				38			

BASIC: For which there is information about children and main woman age, height, gender, and that live in a village where price data was collected. **MAIN:** Same sample as the main regressions

Table 3 shows information on the communities in which these households reside. In total, there are 40 communities that enter the regression analysis[11]. Seventy percent of these communities are slums, the remaining peripheral villages. Similarly as at the household level, close to 50% have access to a toilet. The difference with the household level indicator stems from the fact that statistics presented at the household level zoom in on households with children age 0-5 only whereas the village level one is based on a random sample of households within the community.[12]

A further important variable we include in our analysis is information on water and garbage disposal services. These are factors that are expected to affect the disease environment and their inclusion as conditioning variables is hence important under the presumption that doing so will alleviate omitted variable bias. We construct an indicator that combines information on whether garbage (kitchen and other) is picked up by trucks or disposed in waste baskets, and whether the community receives water for cooking and drinking from a hand-pump or household service connections.[13]

Our main variable of interest is the sanitation environment the study households, and specifically their children, reside in. As discussed in Section 2, we define it as the percentage of households in the village that child i resides in, including its own household, that use a sanitation system for defecation. In the survey, households were asked about the sanitation behaviour of groups of household members (boys, girls, made adults,

[11]For three communities no price information is available, implying that we cannot include them in our main analysis. These three communities do not display any systematic differences to those included though. Information is available upon request.

[12]Households were randomly selected at the time of the baseline survey.

[13]We run regressions with this information included at the disaggregated level as shown in column (7) of Table 6 in the Appendix. The impact coefficient is almost the same, albite slightly noisier. Descriptive statistics on these components are also provided in the Appendix in Table 7.

female adults, male elderly and female elderly). Only if all of these groups report to use the toilet is our indicator equal to one for this household. Usage of community toilets (or usage of neighbour's toilet) is very rare in our sample (as in the Indian context more generally). We however include these households since usage of community toilets of course implies less faeces in the environment. Figure 3 shows a breakdown of usage rates by whether households own their own toilet or not, split by whether they reside in a slum or peripheral village and by survey round. It can be seen that usage rates are very high (85 percent or above) for privately owned toilets, with rates being slightly higher in slums. Households living in slums are also more likely to use toilets (community toilet or that of neighbours) when they do not own one themselves.

Figure 3: Sanitation usage and ownership - by location and round

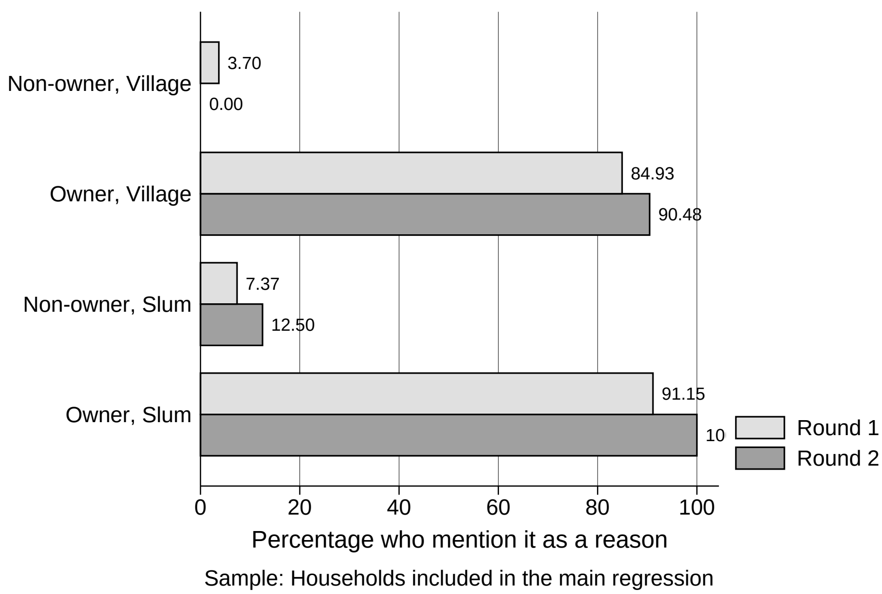
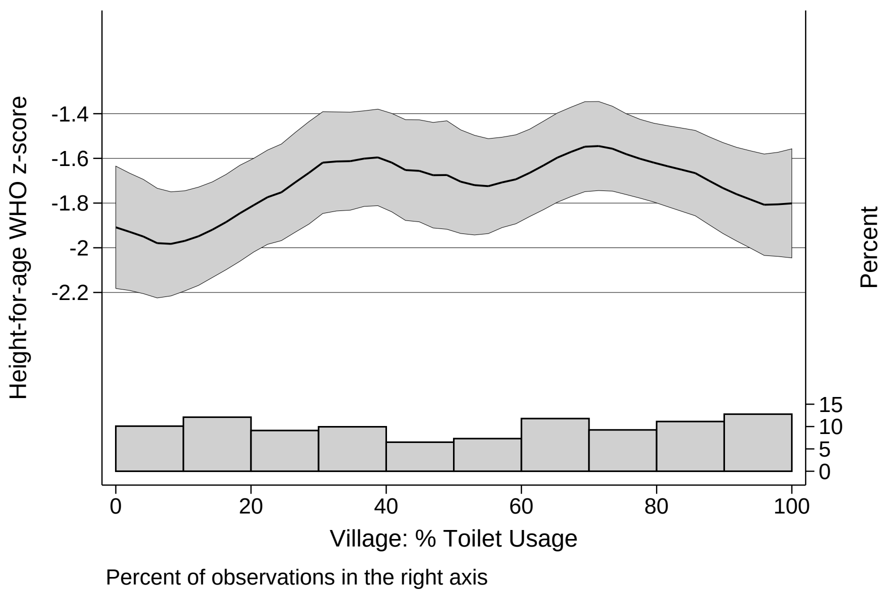


Figure 4 combines our community level sanitation environment measure and the average height z-score for our sample children. The bars at the bottom of this figure illustrate the variation in the construction sanitation environment variable. It can be seen that the sample spans communities where no household uses exclusively sanitation systems when relieving themselves at the one extreme (i.e. $ES_{v,t} = 0$) and at the other extreme communities where in every household every member is reported to use a sanitation system ($ES_{v,t} = 1$). Within these extremes, a wide array of usage fractions are observed in our sample.

The figure further gives a first graphical indication that higher sanitation usage coverage is associated with lower stunting rates. The relationship is not obviously increasing throughout, but seems to increase especially at lower increases of coverage.

We will explore this further in our analysis below.

Figure 4: Height of children and sanitation environment



3.1 Sanitation raw material prices as instruments

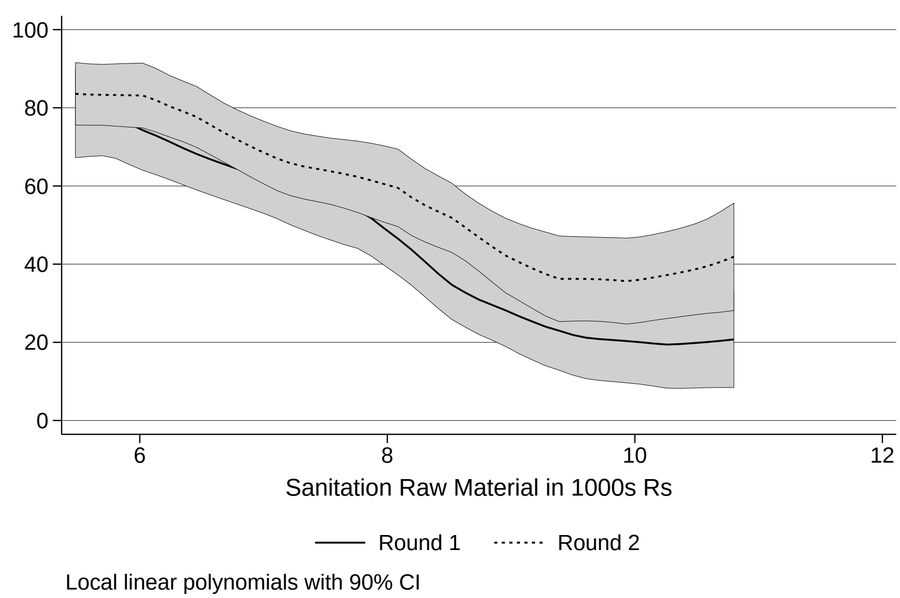
Before discussing our estimation results we need to provide more information on our instrument.

As discussed above, prices are generally acknowledged to affect investment choices without entering the production function directly (Todd and Wolpin [2003], Puentes et al. [2014], Attanasio et al. [2015]) and are hence used, where feasible, to instrument endogenous variables in context of production functions and beyond.

The price data we use was collected shortly after the baseline survey by contacting providers of raw materials for sanitation within the study communities [Gautam, 2015], in line with the requirements of the Government of India's flagship sanitation program (at the time of data collection it was called the Total Sanitation Campaign, recently revamped as the Swachh Bharat Mission).[14] The average raw material prices for a toilet amount to INR 8.300, ranging from about INR 5,500 to 10,800 (see Table 3).

The distribution of prices in relation to sanitation coverage in our communities is displayed in Figure 5, showing, as in previous figures, the distribution of observations across the x-axis in the bottom of the figure. For almost complete sanitation usage coverage, the figure shows a clear downward trend in prices: The higher the price, the lower the coverage of used sanitation infrastructure.

Figure 5: Sanitation raw material prices and sanitation uptake



[14]More details on the collection is provided in Gautam [2015].

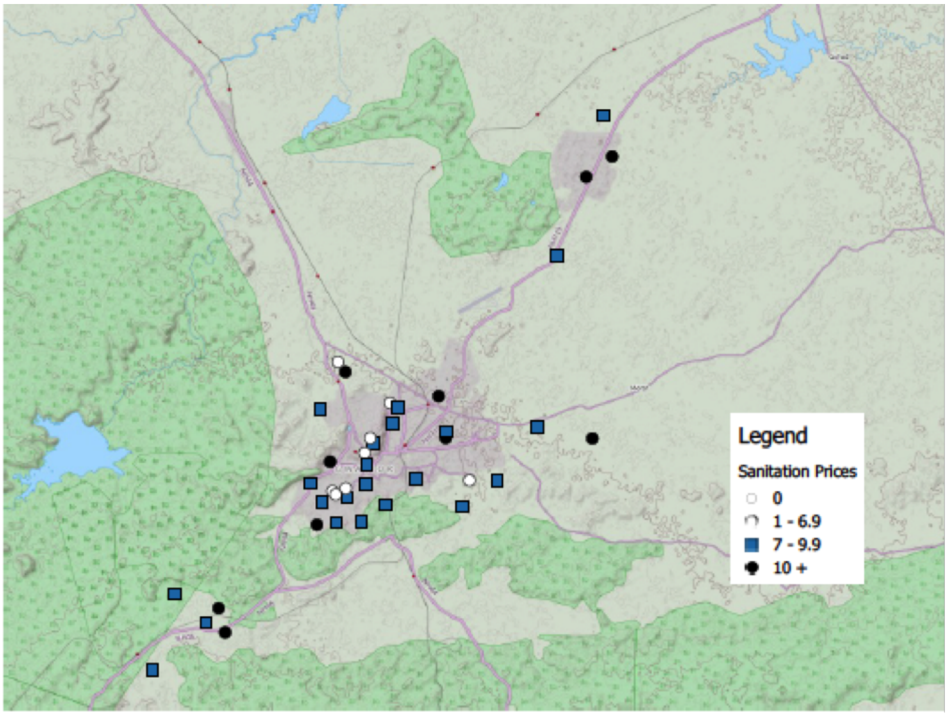
Estimates presented in Figure 5 are the 'pure' first-stage regression, where we only include an indicator of whether the household lives in a slum or peripheral village as well as a survey round indicator, the predictive power of the raw materials is very strong. The F-stat associated with the instrument is 9.73 as shown in the Appendix Table 8, column 3.[15] . We will see in our main regression, that the instrument keeps its strength when included in the full regression, accounting for the complete set of covariates discussed above. Analysing in more detail these prices, we find that they correlate with other price information in the village. This is not an unexpected finding. In line, when including index functions of food prices and access to shops (such as fair price shops, pan shop, tailor, etc.) our estimate becomes less precisely estimated but stays otherwise similar. It suggests that part of the variation in our instrument is related to other prices.

This becomes a concern when the raw material price reflects access to other amenities that might affect health and we do not account for. We can see in the map shown in Figure 6 that the lowest raw material prices are primarily observed within the center of Gwalior (the area with a slightly darker shade of gray). To check whether the prices are a reflection of access to amenities, which could impact child health through other channels, we include information on access to amenities (pharmacy, medical shop, health care center, distance to collectorate and availability of transportation services in our regression. Due to data limitations on these indicators, this exercise is only suggestive as our sample reduces significantly (we are now able to only include 27 clusters in our analysis). We can see however that if anything, our impact estimates become larger and more significant and price differences are more likely driven by supply shocks such as transportation costs.

15

The other columns of the same table present alternative specifications where we control for other village characteristics summarised in three indexes which might be related to both prices and health. Such indexes where constructed by finding variables associated with raw materials prices, and then summarizing them into individual factors using principal factor analysis (Table 7 describe such components). An alternative version of the first stage that includes index components rather than the summary variables is presented in Table 6. Given that the indexes were constructed to explain the variation on toilet prices (see Table 9), it is expected that the inclusion of all index at the same time will reduce the association between the instrument and health.

Figure 6: Community locations and prices of sanitation raw materials



4 Results

We now turn to presenting and discussing our main findings as well as some additional checks we conduct.

Our main finding is shown in Table 4.[16] Column (1) shows the OLS regression, which does not take into account the endogeneity of the sanitation environment we consider in our study. The regression results of this specification suggest that the sanitation usage density is not significantly related to child height for age. Once accounting for the endogeneity though through instrumenting with prices of raw materials, we see in column (2) that the coefficient becomes larger and is now significant at the ten percent level. The fact that the OLS estimate is downward biased can be understood in the context of the example given in Section 2.1, where one can imagine village with worse condition for health, such as very high population density coupled with limited (health) infrastructure to be more likely to make investments in private household sanitation, which improves the wider sanitation and hence disease environment - leading to a negative, or lower, correlation between sanitation environment and health.

[16]Full regression results with information on all covariates are shown in Table 10 in the appendix. .

Table 4: Avg Sanitation and height-for-age

Sample: Children aged 0 to 5.		
	(1) OLS	(2) IV
Panel A: Second Stage		
Village % who uses a toilet	0.004	0.017**
	(0.005)	(0.008)
Girl	0.040	0.083
	(0.125)	(0.124)
Panel B: First Stage		
Sanitation Raw Mat Price (1000 Rps)		−8.053***
		(2.242)
F-Stat		12.90
Obs	891	863
Clust	41	40
R2 Adj	0.11	0.10

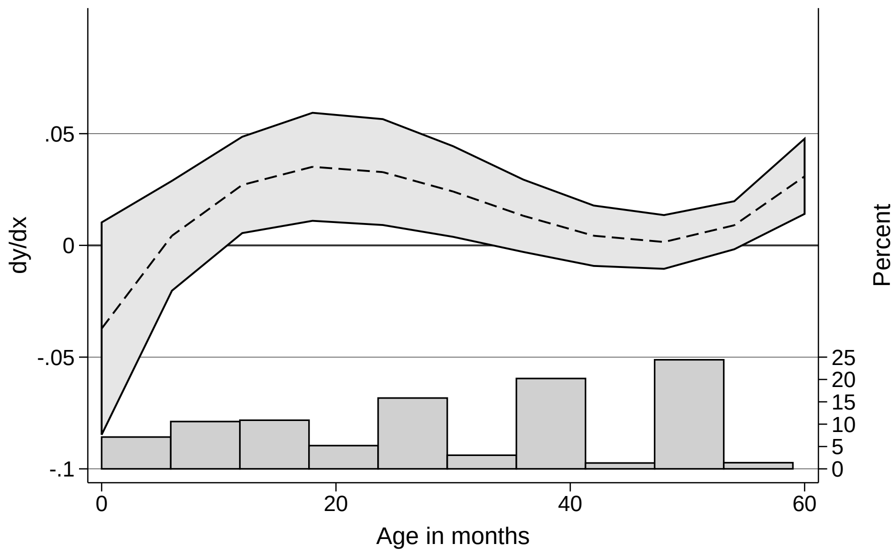
Sample: Children aged 0 to 5.
Controls: 3rd order polynomial on age, gender, mother education, cuartiles of income, HH size, N of HH memebers who are males, any adverse shock last 2 years, slum and wave dummies, and a factor for quality of water and waste deposition. SE clustered at Village level in parenthesis. Significance: * 10%, ** 5%, *** 1%.

The F-stat of the IV regression is 12.89, observably lower than in the 'pure' regression due to accounting for a large set of covariates at the child, household and community level but still strong.

Figure 7 shows the marginal effects by age of the child and we can see that the effects are positive and increasing in the approximate age range of six to around 22 months of age. This observation is in line with the idea that a hygienic environment and breastfeeding can serve as substitutes due to antibodies in the milk (Van der Slice, Popkin and Briscoe 1994). Given that breastfeeding is typically advised to be done exclusively for the first six months of life we would not expect an improvement in sanitation to have as much of an effect. On the other hand, as previously mentioned, most placidity in growth is seen in the first two years of life (Victoria et al. [2010]), providing hence the largest opportunity for improvements in the sanitation environment to have impacts on child height for age.

The coefficient of 0.017 suggests that a 10% increase in sanitation usage coverage increases child height on average by 0.17 standard deviations of the z-score. To put this number in context, an increase of ten percentage point in sanitation coverage is translated into approximately 0.7 centimetres increase for a four year old child. We show in Figure 8 which translates the estimated impacts into centimetres at all ages. The dashed line ("Homog. joint") are the estimated effects for the full sample, the solid line

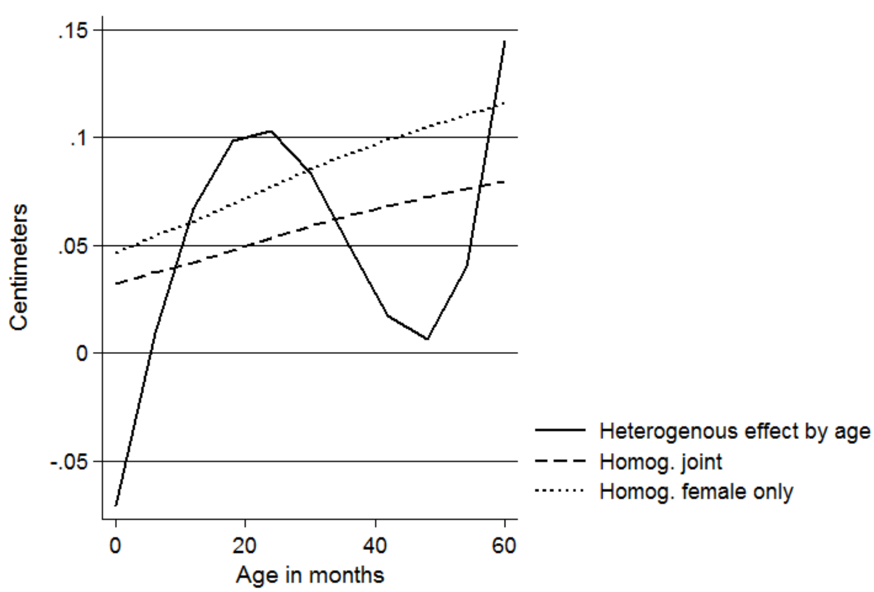
Figure 7: Marginal effects by age



CI at 90% level based on SE clustered at Village Level. Marginal effects after 2SLS.

shows impacts by age, reflecting the importance of the first two years in life.

Figure 8: Translation of estimated effects into centimetres



This compares to the estimated impact of the working paper by Hammer [2013]: Their estimates suggest a program impact of 0.3-0.4 standard deviations, which translates into an increase in toilet ownership of 8.2 percentage points, which they postulate impacts child height by 1.3 centimetres in four year old children. This impact shown in Hammer

[2013] has been suggested to be larger than biologically plausible when compared to for example to findings by Richard et al. [2013]. In this study, the authors use data from seven cohort studies to understand the lagged association between diarrhoea and growth in the first two years of life, a period seen as the greatest plasticity in growth trajectory (Victoria et al. [2010]). They find the cumulative effect on child's length from diarrhoea burden in these first two years of life to be 0.38 centimetres .

When splitting these impacts by gender of the child we find that they are primarily driven by impacts on female children.

Table 5 shows these findings. As in Table 4, the first columns are findings from the OLS regressions, the right panel from the IV ones. We show for each set of estimation approaches first sub-sample regressions for males (columns 1 and 4) and females (columns 2 and 5). Columns 3 and 6 include an interaction of our sanitation environment variable and the gender of the child. Similarly, the instrument is interacted with the information on gender. We find that the impact on boys is insignificant while that for girls is highly significant at one percent and also the instruments exhibit greater power in this specification.

Table 5: Avg Sanitation and height-for-age

Sample: Children aged 0 to 5.						
	(1)	(2) OLS	(3)	(4)	(5) IV	(6)
	MALE	FEMALE	BOTH	MALE	FEMALE	BOTH
Panel A: Second Stage						
Village % who uses a toilet	0.002	0.007		0.008	0.025***	
	(0.006)	(0.005)		(0.009)	(0.009)	
Village Avg * Boy			0.003			0.014
			(0.005)			(0.009)
Village Avg * Girl			0.004			0.021***
			(0.005)			(0.008)
Girl			−0.000			−0.235
			(0.200)			(0.291)
Panel B: First Stage						
Sanitation Raw Mat Price (1000 Rps)				−8.247***	−8.045***	
				(2.110)	(2.309)	
F-Stat				15.28	12.13	18.61/ 12.93
Obs	458	433	891	445	418	863
Clust	40	40	41	39	39	40
R2 Adj	0.07	0.15	0.11	0.07	0.12	0.10
H0: $\beta_{Girls} - \beta_{Boys} = 0$			0.84			0.25

Controls: 3rd order polynomial on age, gender, mother education, cuartiles of income, HH size, N of HH memebers who are males, any adverse shock last 2 years, slum and wave dummies, and a factor for quality of water and waste deposition. SE clustered at Village level in parenthesis. Significance: * 10%, ** 5%, *** 1%.

4.1 Possible mechanisms of differential gender impacts

It is not immediately obvious why girls should benefit more from an improvement in the sanitation environment they live in. We bring forward two reasons we think might explain this finding. Unfortunately, our data does not allow us to formally test for the relevance of these and are hence only suggestive.

The first possible mechanism that could explain the differential impacts by gender relate to preferential investment. It is known that Indian families have explicit preferences for having sons over daughters [Pande and Astone, 2007]. This male preference translates into differential investment: evidence suggests that "[...] boys receive more childcare time than girls, they are breastfed longer and they get more vitamin supplementation" [Carvalho et al., 2013]. There is further evidence that boys receive more nutrition (Das Gupta [1987]), more healthcare (Basu [1989], Ganatra and Hirve [1994]), are breastfed for longer [Jayachandran and Kuziemko, 2015], and are more likely to be vaccinated [Borooah, 2004] - all investments that are known to boost the immune system. Breastfeeding for example is shown to provide important inputs for the immune system, which some argue can act as a substitute for sanitation [VanDerslice et al., 1994].

Such differential investments in important inputs for the immune system could hence lead girls to respond more positively to improvements in the sanitation environment.

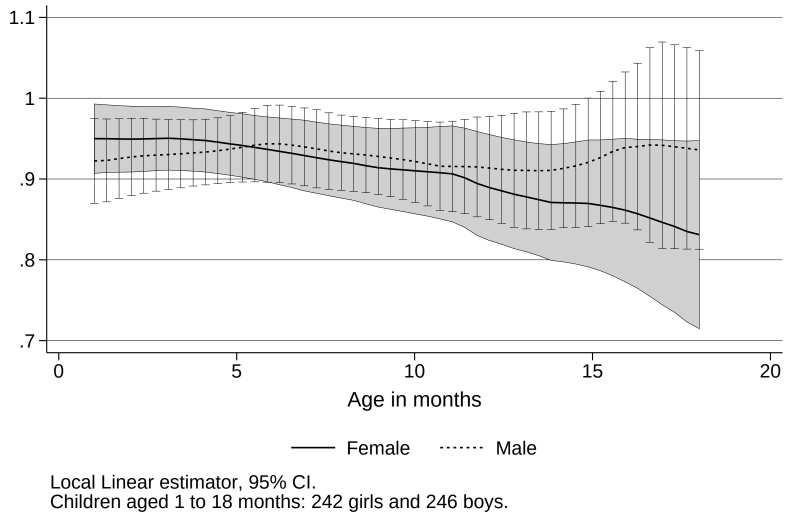
As discussed already in Section 2.3, our data is faces strong limitation in terms of what and how well nutritional inputs, including breastfeeding, are measured. With this caveat in mind, our data suggests that the differential investment might not be an important channel behind the gender findings of our analysis. While we see in Figure 9 that breastfeeding stops earlier for girls in our sample than for boys,[17] confirming other studies on boys preferences, this descriptive difference is not significant. We also compared nutritional intake for children aged 2 to 5, but we found no evidence of a systematic gender difference[18].

The second possible mechanisms behind the differential impacts by gender found in our analysis relates to the possibility of girls being more directly affected than boys by the increase in sanitation usage coverage. Our data shows that if a toilet owned is not used by all household members, it is boys and men who are least likely to use the toilet. This suggests that girls frequent open defecation sites less than boys (possibly since their mums do not go anymore and hence do not take the girls along), and are hence

[17]The graphs shows the information for our sub-sample of children age 0-18 months for which information on breastfeeding is available.

[18]More details in section A.3 in the Appendix.

Figure 9: Breastfeeding by age and gender



exposed to a cleaner environment than boys (despite boys frequenting open defecating areas used by less people no average). Data limitations unfortunately stop us from analysing this possible mechanism in more detail.

5 Conclusion

We make use of primary data collected as part of an evaluation exercise of a sanitation intervention to investigate the impact of improvements in the sanitation environment, defined as the fraction of households using private or community toilets, on child height for age, an indicator for health.

We do so in the context of slums and peripheral villages in a city in Northern India, Gwalior. This population is an important one to consider for two main reasons: India's slum population is growing rapidly while at the same time having no or only inadequate access to safe sanitation. High population density coupled with improper means of disposing faeces provides a breeding ground for preventable disease epidemics. Providing evidence on improvements in children's health that can be achieved by community-level sanitation improvements is hence of direct policy relevance.

Our results suggest that increases in sanitation usage rates significantly affect children's height. This impact seems to be particularly relevant for girls. We suggest two possible mechanisms behind these impacts. Unfortunately we are only able to provide suggestive evidence for these mechanisms in our study population given data and sample size restrictions. However, independent of the drivers behind this finding, the results suggest that not only is investment in sanitation coverage worthwhile when children's health is one of the objectives, but increasing sanitation coverage seems to be at the same time a policy that implicitly targets girls.

References

S.C. Aggarwal and N. Kumar. Level of poverty and employment pattern in slums: A case of gwalior in central india. *The Indian Journal of Labour Economics*, 51(2):315–30, 2008. 2.2

D. Almond and J. Currie. *Human Capital Development Before Age 5*, volume pp.1315-1486. Elsevier, 2011. 1

L.A. Andres, B. Briceño, C. Chase, , and J.A. Echenique. Sanitation and externalities: Evidence from early childhood health in rural india. Technical report, World Bank, 2014. http://dx.doi.org/10.1596/1813-9450-6737. 1, 2.1

O. Attanasio, S. Cattan, Fitzsimons, C. Meghir, and M. Rubio-Codina. Estimating the production function for human capital: Results from a randomized control trial in colombia. NBER Working Paper No. 20965, February 2015. 1, 2.2, 2.2

A.M. Basu. Is discrimination in food really necessary for explaining sex differentials in childhood mortality. *Population Studies*, 43:193–210, 1989. 4.1

J. Behrman, M. Calderon, S.H. Preston, J.F. Hoddinott, R. Martorell, and A.D. Stein. Nutritional supplementation of girls influences the growth of their children: Prospective study in guatemala. *American Journal of Clinical Nutrition*, 90(5):1372–9, 2009. 1

R. E. Black, L. H. Allen, Z. A. Bhutta, L. E. Caulfield, M. de Onis, M. Ezzati, C. Mathers, and J. Rivera. Maternal and child undernutrition: Global and regional exposures and health consequences. *The Lancet*, 371(9608):243–60, 2008. 1

Vani K Borooah. Gender bias among children in india in their diet and immunisation against disease. *Social Science & Medicine*, 58(9):1719–1731, 2004. 4.1

C.G. Bozzoli, A.S. Deaton, and C. Quintana-Domeque. Adult height and childhood disease. *Demography*, 46(4):647–69, 2009. 1

Leandro S Carvalho, Silvia H Barcellos, and Adriana Lleras-Muney. Child gender and parental investments in india: Are boys and girls treated differently? *American Economic Journal: Applied Economics*, 2013. 4.1

W. Checkley, G. Buckley, R.H. Gilman, A.M. Assis, R. Guerrant, S.S. Morris, K. Molbak, P. Valentiner-Branth, C.F. Lanata, and R.E. Black. Multi-country analysis of the effects of diarrhoea on childhood stunting. *International Journal of Epidemiology*, 37(4):816–30, 2008. 1

Thomas Clasen, Sophie Boisson, Parimita Routray, Belen Torondel, Melissa Bell, Oliver Cumming, Jeroen Ensink, Matthew Freeman, Marion Jenkins, Mitsunori Odagiri, et al. Effectiveness of a rural sanitation programme on diarrhoea, soil-transmitted helminth infection, and child malnutrition in odisha, india: a cluster-randomised trial. *The Lancet Global Health*, 2014. 1, 3

J. Currie. *Handbook of Health Economics*, volume 1, chapter 19. Child Health in developed countries, pages 1053–1090. Elsevier, 2000. 2.1

M. Das Gupta. Selective discrimination against female children in rural punjab, india. *Population and Development Review*, 13:377–400, 1987. 4.1

Mercedes de Onis, David Brown, Monika Blossner, and Elaine Borghi. Levels and trends in child malnutrition. joint child malnutrition estimates. Technical report, UNICEF Geneva and WHO Washington DC and World Bank, 2015. 1

A. Deaton. Height, health, and development. *Proceedings of the National Academy of Sciences*, 104 (33):13232–13237, 2007. 1

E. Duflo, M. Greenstone, R. Guiteras, and T. Clasen. Toilets can work - short and medium run health impacts of addressing complementarities and externalities in water and sanitation. NBER Working Paper w21521, 2015a. 2.1

Esther Duflo, Michael Greenstone, Raymond Guiteras, and Thomas Clasen. Toilets can work: Short and medium run health impacts of addressing complementarities and externalities in water and sanitation. Technical report, National Bureau of Economic Research, 2015b. 1

B. Ganatra and S. Hirve. Male bias in health care utilization for under-fives in a rural community in western india. *World Health Organization Bulletin*, 72(1):101–4, 1994. 4.1

Sanghmitra Gautam. Quantifying welfare effects in the presence of externalites: An ex-ante evaluation of a sanitation intervention. Unpublished, 2015. 2.2, 14

P. Gertler, M. Shah, L. Cameron, S. Martinez, and S. Patil. How does health promotion work? evidence from the dirty business of eliminatingopen defecation. December 2014. 1, 2.1

M. Geruso and D. Spears. Sanitation and health externalities: Resolving the muslim mortality paradox. January 2014. 1, 2.1, 10

E. Ginsburg, G. Livshits, K. Yakovenko, and E. Kobyliansky. Major gene control of human body height, weight and bmi in five ethnically different populations. *Annals of Human Genetics*, 62: 307–322, 1998. 2.2

Dean Hammer, Jeffrey; Spears. Village sanitation and children's human capital : Evidence from a randomized experiment by the maharashtra government. Technical report, World Bank, 2013. URL https://openknowledge.worldbank.org/handle/10986/16014. License: CC BY 3.0 Unported. 4

P. Hathi, S. Haque, L. Pant, D. Coffey, and D. Spears. Place and child health - the interaction between density and sanitation in developing countries. Technical report, Policy Research Working Paper 7124 World Bank Group, November 2014. 2.2

J. Hoddinott, J. Behrman, J.A. Maluccio, P. Melgar, A.R. Quisumbing, M. Ramirez-Zea, A.D. Stein, K.M. Yount, and R. Martorell. Adult consequences of growth failure in early childhood. *The American Journal of Clinical Nutrition*, 98(5):1170–78, 2013. 1

J. F. Hoddinott, J. A. Maluccio, J. R. Behrman, R. Flores, and R. Martorell. Effect of a nutrition intervention during early childhood on economic productivity in guatemalan adults. *The Lancet*, 371(9610):411–6, 2008. 1

S Jayachandran and I Kuziemko. Why do mothers breastfeed girls less than boys? evidence and implications for child health in india. *Quarterly Journal of Economics*, 126(3):1485–1538, 2015. 4.1

J. A. Maluccio, J. F. Hoddinott, J. R. Behrman, A. R. Quisumbing, R. Martorell, and A. D. Stein. The impact of nutrition during early childhood on education among guatemalan adults. *Economic Journal*, 119(537):734–63, 2009. 1

W.H. Mosley and L.C. Chen. An analytical framework for the study of child survival in developing countries. *Bulletin of the World Health Organization*, 81(2):140–145, 1984. 2.1

C. Nokes, S.M. Grantham-McGregor, A.W. Sawyer, E.S. Cooper, B.A. Rpbinson, and D.A.P. Bundy. Moderate-to-heavy infection of trichuris trichiura affect cognitive function in jamaican school children. *Parasitology*, 104:539–547, 1992. 1

C. Nokes, C. van den Bosch, and D.A.P. Bundy. The effects of iron deficiency and anemia on mental and motor performance, educational achievement, and behavior in children: A report of the international nutritional anemia consultative group. Technical report, USAID: Washington, DC, 1998. 1

NSSO. Some characteristics of urban sslum 2008-09. Report 534, National Sample Survey Office, National Statistical Organization, Ministry of Statistics and Programme Implementation, Government of India, May 2010. 2.2

Rohini P Pande and Nan Marie Astone. Explaining son preference in rural india: the independent role of structural versus individual factors. *Population Research and Policy Review*, 26(1):1–29, 2007. 4.1

A.J. Pickering, H. Djebbari, C. Lopez, M. Coulibaly, and M.L. Alzua. Effect of a community-led sanitation intervention on child diarrhoea and child growth in rural mali: a cluster-randomised controlled trial. *The Lancet*, 3(11):701–711, 2015. 2

Annette Pruesss-Uestuen, Robert Bos, Fiona Gore, Jamie Bartram, et al. *Safer water, better health: costs, benefits and sustainability of interventions to protect and promote health.* World Health Organization, 2008. 1

E. Puentes, F. Wang, J.R. Behrman, F. Cunha, J. Hoddinott, J.A. Maluccio, L.S. Adair, J.B. Borja, R. Martorell, and A.D. Stei. Early life height and weight production functions with endogenous energy and protein inputs. September 2014. 1, 2.2, 2.2, 2.2, A.3

S.A. Richard, R.E. Black, R.H. Gilman, R.L. Guerrant, G. Kang, C.F. Lanata, K. Mølbak, Z.A. Rasmussen, R.B. Sack, P. Valentiner-Branth, and W. Checkley. Diarrhea in early childhood: Short-term association with weight and long-term association with length. *American Journal of Epidemiolog*, 178(7), August 2013. doi: 10.1093/aje/kwt094. 4

K. Silventoinen. Determinants of variation in adult body height. *Journal of Biosocial Science*, 35(2): 263–285, 2003. 2.2

L Smith and L. Haddad. Reducing child undernutrition: Past drivers and priorities for the post-mdg era. Institute of Development Studies Working Paper. Volume 2014 No. 441., April 2014. 1

P.E. Todd and K.I. Wolpin. On the specification and estimation of the production function for cognitive achievement. *The Economic Journal*, 113(485):F3–F33, 2003. 1, 2.2, 6, 2.2

James VanDerslice, B Popkin, and John Briscoe. Drinking-water quality, sanitation, and breast-feeding: their interactive effects on infant health. *Bulletin of the World health organization*, 72(4): 589, 1994. 4.1

C.G. Victora, M. de Onis, P.C. Hallal, M. Bloessner, and R. Shrimpton. Worldwide timing of growth faltering: Revisiting implications for interventions. *Pediatrics*, 125(3):e473–e480, 2010. 1

C.G. Victoria, M. de Onis, P.C. Hallal, M. Bloessner, and R. Shrimpton. Worldwide timing of growth faltering: Revisiting implications for interventions. *Pediatrics*, 125(3):e473–480, 2010. 2.2, 4, 4

S. P. Walker, T. D Wachs, S. Grantham-McGregor, M. M. Black, C. A Nelson, S. L. Huffman, H. Baker-Henningham, S. M. Chang, J. D. Hamadani, B. Lozoff, J. M. Meeks Gardner, C. A. Powell, A. Rahman, and L. Richter. Inequality in early childhood: Risk and protective factors for early child development. *The Lancet*, 378(9799):1325–38, 2011. 1

World Health Organization WHO. Poor sanitation threatens public health, March 2008. URL http://www.who.int/mediacentre/news/releases/2008/pr08/en/. 1

World Health Organization. Fact sheet n 330: Diarrhoeal disease, April 2013. URL http://www.who.int/mediacentre/factsheets/fs330/en/. 1

A Appendix

A.1 'Pure' first stage regression

Table 6: First Stage and components of the indexes (Round 1 only)

Dependent: % of HH whose members use a toilet	(1)	(2)	(3)	(4)
Cost of Raw Materials (1000s Rps)	−7.939*	−8.940**	−8.520***	−7.084
	(4.209)	(4.000)	(2.915)	(4.809)
Were new dwellings built in this village in the last 12 months?	29.957			33.409
	(17.989)			(30.705)
Are autos available to drive to this bus stop?	4.138			12.692
	(10.111)			(11.344)
Village has kirana/general market shop?	−51.000**			−43.580*
	(24.212)			(23.714)
Village has wine shop?	−4.012			−10.315
	(13.991)			(16.733)
Village has tailoring shop?	4.430			3.224
	(13.797)			(13.500)
Village has fair price shop?	24.631			14.374
	(21.669)			(23.164)
Village has paan shop?	11.093			7.838
	(9.181)			(9.702)
Village has mahila mandal?	−29.621			−21.226
	(19.002)			(22.408)
Village has community centre?	−19.765			−23.573
	(13.559)			(23.685)
Village has library?	11.883			13.873
	(15.058)			(26.751)
Village has panchayat office?	4.033			7.256
	(9.965)			(13.860)
Village has fair price shop?	−18.573			−14.873
	(17.906)			(20.419)
Village has playground?	−12.124			−7.439
	(14.404)			(11.471)
Price of 1kg sugar from market		−0.307		1.852
		(1.156)		(1.627)
Price of 1l edible oils		0.168		−0.772
		(0.421)		(0.789)
Price of 1kg onions		−1.166		0.637
		(1.224)		(2.014)
Price of 1kg chicken		0.133		−0.272
		(0.157)		(0.319)
Price of 1 tea		0.116		−0.008
		(0.071)		(0.119)
Throw kitchen garbage away in waste baskets/trucks pick it up?			38.648***	40.397
			(9.087)	(25.615)
Does the village community get water for cooking and drinking from hand pump?			−5.616	−11.679
			(7.731)	(19.713)
Does the village get water for cooking/drinking from household service connectio			−0.216	−7.601
			(8.943)	(18.924)
Village/slum is a slum	25.339***	24.476*	9.585	10.726
	(7.861)	(12.680)	(7.895)	(11.559)
N Observations	37	39	40	36
R2 Adj	0.65	0.49	0.67	0.64
F-stat instrument	3.56	5.00	8.54	2.17
DF	36	38	39	35
p-val	0.0674	0.0314	0.0058	0.1497

SE clustered at village level in parenthesis. Significance level: * 10%, ** 5%, *** 1%

Table 7: Components of the Indexes (Round 1)

Variable	Mean	Std Dev	Correl. with Price Raw Materials	Correl. with its Index
	(1)	(2)	(3)	
Village Scale and Location Index				
Were new dwellings built in this village in the last 12 months?	0.075	0.267	−0.297***	0.587***
Are autos available to drive to this bus stop?	0.875	0.335	−0.323***	0.223***
Village has kirana/general market shop?	0.951	0.218	−0.269***	0.251***
Village has wine shop?	0.400	0.496	−0.289***	0.796***
Village has tailoring shop?	0.750	0.439	−0.396***	0.582***
Village has fair price shop?	0.500	0.506	−0.320***	0.786***
Village has paan shop?	0.475	0.506	−0.316***	0.670***
Village has mahila mandal?	0.150	0.362	0.324***	0.115***
Village has community centre?	0.175	0.385	−0.272***	0.627***
Village has library?	0.050	0.221	−0.302***	0.421***
Village has panchayat office?	0.250	0.439	0.315***	0.186***
Village has fair price shop?	0.350	0.483	−0.305***	0.578***
Village has playground?	0.350	0.483	0.292***	0.357***
General Prices Index				
Price of 1kg sugar from market	40.200	3.757	−0.358***	0.483***
Price of 1l edible oils	55.050	6.664	0.333***	−0.465***
Price of 1kg onions	14.925	4.015	−0.380***	0.427***
Price of 1kg chicken	101.250	30.900	−0.508***	0.899***
Price of 1 tea	30.175	61.096	−0.522***	0.877***
Water and Garbage disposal Index				
Throw kitchen garbage away in waste baskets/trucks pick it up?	0.293	0.461	−0.328***	0.887***
Does the village community get water for cooking and drinking from hand pump?	0.537	0.505	0.433***	−0.249***
Does the village get water for cooking/drinking from household service connectio	0.390	0.494	−0.317***	0.945***

The first column is the average of the variable. The second corresponds to the correlation with raw material prices. The last one presents the correlation with the specific index.
Significance level: * 10%, ** 5%, *** 1%

Table 8: First stage regression, prices and limited covariates on health

Dependent: % of HH whose members use a toilet	Round 1	Round 2			Both rounds †		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Cost of Raw Materials (1000s Rps)	−5.668*	−11.297***	−8.679***	−3.997	−8.354***	−5.655	−6.577**
	(3.018)	(3.018)	(2.783)	(3.250)	(2.842)	(3.655)	(2.434)
Village/slum is a slum	34.279***	22.508**	27.809***	18.239*	31.469***	29.105***	15.281**
	(9.313)	(8.739)	(8.162)	(9.179)	(8.933)	(9.368)	(7.458)
Baseline Observation			−12.198***	−12.232***	−11.223***	−12.091***	−12.587***
			(3.159)	(3.646)	(3.412)	(3.315)	(3.323)
Village Scale and Location Index				1.530	2.509		
				(4.309)	(4.511)		
General Prices Index				5.414		7.942*	
				(5.122)		(4.456)	
Water and Garbage disposal Index				17.086***			18.064***
				(5.824)			(4.629)
N Observations	38	40	78	68	69	74	75
N Villages	38	40	43	36	37	39	40
R2 Adj	0.46	0.49	0.49	0.63	0.52	0.51	0.63
F-stat instrument	3.53	14.02	9.72	1.51	8.64	2.39	7.30
DF	37	39	42	35	36	38	39
p-val	0.0682	0.0006	0.0033	0.2270	0.0057	0.1301	0.0101

† SE clustered at village level in parenthesis. Significance level: * 10%, ** 5%, *** 1%

Table 9: Raw Materials Prices and Village Indexes

	(1)	(2)	(3)	(4)
Village Scale and Location Index	-0.488**	-0.760***		
	(0.234)	(0.252)		
General Prices Index	-0.931***		-1.013***	
	(0.161)		(0.162)	
Water and Garbage disposal Index	-0.321			-0.522
	(0.274)			(0.352)
Village/slum is a slum	-0.600	-1.423***	-1.033**	-1.020*
	(0.529)	(0.511)	(0.449)	(0.574)
Baseline Observation	0.0191	-0.0300	-0.0195	-0.0778
	(0.0544)	(0.0721)	(0.0548)	(0.104)
N Observations	68	69	74	75
N Villages	36	37	39	40
R Sqrd	0.556	0.297	0.459	0.207

SE clustered at village level in parenthesis. Significance level: * 10%, ** 5%, *** 1%

A.2 Full regression results - main specification

Table 10: Avg Sanitation and height-for-age (Detailed)

Sample: Children aged 0 to 5.			
	(1) OLS	(2) IV-FS	(3) IV-SS
Panel A: Second Stage			
Village % who uses a toilet	0.004		0.017**
	(0.005)		(0.008)
Cost of Raw Materials (1000s Rps)		−8.053***	
		(2.242)	
Girl	0.040	0.787	0.083
	(0.125)	(1.062)	(0.124)
Age in months	−0.203***	0.362	−0.212***
	(0.055)	(0.335)	(0.055)
Age in months ^{2} (100s)	0.606***	−1.447	0.636***
	(0.192)	(1.279)	(0.193)
Age in months ^{3} (10000s)	−0.513**	1.557	−0.543***
	(0.190)	(1.317)	(0.190)
Mother Height (cm)	0.045***	−0.258*	0.048***
	(0.012)	(0.129)	(0.012)
Mother age	0.007	0.050	0.007
	(0.008)	(0.072)	(0.008)
Village/slum is a slum	−0.406***	12.730	−0.689**
	(0.148)	(7.623)	(0.273)
Baseline Observation	−0.966***	−13.013***	−0.861***
	(0.277)	(2.964)	(0.293)
Mother Education: 6-8 yrs	0.081	2.632	0.007
	(0.266)	(2.445)	(0.278)
Mother Education: 9 yrs +	0.047	0.570	−0.031
	(0.257)	(2.101)	(0.263)
HH self-reported Income	−0.000	0.002	−0.000
	(0.003)	(0.046)	(0.003)
Total consumption expenditures of hh in last year	0.001	0.016*	0.001
	(0.001)	(0.009)	(0.001)
Income: 40.15 - 60.00 K Rup	−0.055	−2.660	0.003
	(0.182)	(2.480)	(0.189)
Income: 60.25 - 90.00 K Rup	0.231	−1.567	0.187
	(0.250)	(3.749)	(0.249)
Income: 90.30 - 280.00 K Rup	0.507	−3.297	0.500
	(0.398)	(6.945)	(0.415)
Type of dwelling: strong	−0.098	5.967	−0.207
	(0.256)	(4.790)	(0.283)
Type of dwelling: semi-strong	−0.282	−0.526	−0.266
	(0.218)	(3.926)	(0.227)
Number of HH members	−0.110**	0.057	−0.130***
	(0.048)	(0.373)	(0.048)
Number of male HH members	0.025	0.171	0.045
	(0.052)	(0.770)	(0.053)
Any household shock last 12 months	−0.205	−4.227	−0.102
	(0.339)	(3.420)	(0.345)
=1 if Muslim	0.153	−5.484	0.112
	(0.214)	(3.814)	(0.236)

Continued on next page

Table 10: (Continued)

Sample: Children aged 0 to 5.			
	(1)	(2)	(2)
	OLS	IV-FS	IV-SS
=1 if unkown religion	1.168	7.702	0.925
	(1.968)	(4.737)	(1.802)
=1 if forward caste	−0.229	5.640*	−0.312
	(0.273)	(2.904)	(0.309)
=1 if minority backward caste	−0.037	−1.860	−0.010
	(0.419)	(2.973)	(0.450)
=1 if scheduled caste or tribe	−0.224	1.439	−0.317
	(0.240)	(2.420)	(0.246)
=1 if unkown Caste	0.442	1.389	0.391
	(0.428)	(2.847)	(0.424)
Water and Garbage disposal Index	−0.160	16.853***	−0.445**
	(0.167)	(3.906)	(0.182)
Panel B: First Stage			
F-Stat			12.90
Obs	891	863	863
Clust	41	40	40
R2 Adj	0.11	0.73	0.10

SE clustered at Village level in parenthesis. Significance: * 10%, ** 5%, *** 1%.

Table 11: Avg Sanitation and height-for-age (Sets of Covariates)

Sample: Children aged 0 to 5.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Second Stage							
Village % who uses a toilet	0.017**	0.012**	0.014**	0.012*	0.015*	0.028	0.019*
	(0.008)	(0.006)	(0.007)	(0.006)	(0.008)	(0.035)	(0.011)
Girl	0.083	0.002	−0.009	−0.013	0.035	−0.007	0.034
	(0.124)	(0.114)	(0.123)	(0.126)	(0.132)	(0.150)	(0.240)
Age in months	−0.212***	−0.168***	−0.169***	−0.166***	−0.234***	−0.215***	−0.246***
	(0.055)	(0.059)	(0.058)	(0.059)	(0.052)	(0.055)	(0.091)
Age in months ^{2} (100s)	0.636***	0.505***	0.505***	0.500**	0.708***	0.656***	0.637**
	(0.193)	(0.196)	(0.194)	(0.197)	(0.181)	(0.191)	(0.313)
Age in months ^{3} (10000s)	−0.543***	−0.428**	−0.428**	−0.425**	−0.610***	−0.574***	−0.467
	(0.190)	(0.186)	(0.185)	(0.188)	(0.178)	(0.187)	(0.299)
Mother Height (cm)	0.048***	0.042***	0.042***	0.040***	0.049***	0.056***	0.049**
	(0.012)	(0.012)	(0.012)	(0.013)	(0.013)	(0.017)	(0.023)
Mother age	0.007	0.002	0.005	0.006	0.007	0.001	0.011
	(0.008)	(0.008)	(0.009)	(0.008)	(0.008)	(0.010)	(0.013)
Village/slum is a slum	−0.689**	−0.747***	−0.784***	−0.697**	−0.957***	−1.272	−0.760**
	(0.273)	(0.253)	(0.287)	(0.271)	(0.347)	(1.247)	(0.352)
Baseline Observation	−0.861***	−0.871***	−0.859***	−0.879***	−0.851***	−0.770	0.000
	(0.293)	(0.267)	(0.304)	(0.298)	(0.320)	(0.682)	(0.000)
Mother Education: 6-8 yrs	0.007		0.059	0.091	−0.017	−0.190	0.619
	(0.278)		(0.277)	(0.264)	(0.299)	(0.341)	(0.416)
Mother Education: 9 yrs +	−0.031		−0.053	−0.016	−0.149	−0.129	0.117
	(0.263)		(0.254)	(0.264)	(0.299)	(0.368)	(0.367)
HH self-reported Income	−0.000		0.000	0.000	−0.000	−0.001	0.001
	(0.003)		(0.003)	(0.003)	(0.003)	(0.004)	(0.007)
Total consumption expenditures of hh in last year	0.001		0.000	−0.000	0.000	0.000	0.003
	(0.001)		(0.001)	(0.001)	(0.001)	(0.001)	(0.004)
Income: 40.15 - 60.00 K Rup	0.003		−0.003	−0.020	−0.050	0.040	0.221
	(0.189)		(0.174)	(0.176)	(0.191)	(0.209)	(0.345)
Income: 60.25 - 90.00 K Rup	0.187		0.086	0.096	0.234	0.151	0.282
	(0.249)		(0.255)	(0.248)	(0.243)	(0.287)	(0.440)
Income: 90.30 - 280.00 K Rup	0.500		0.472	0.481	0.544	0.575	0.473
	(0.415)		(0.428)	(0.418)	(0.410)	(0.504)	(0.772)
Type of dwelling: strong	−0.207		−0.269	−0.216	−0.161	−0.340	−0.267
	(0.283)		(0.301)	(0.296)	(0.284)	(0.376)	(0.447)
Type of dwelling: semi-strong	−0.266		−0.235	−0.207	−0.100	−0.201	−0.112
	(0.227)		(0.221)	(0.221)	(0.199)	(0.250)	(0.417)
Number of HH members	−0.130***		−0.098**	−0.102**	−0.110**	−0.090*	−0.337**
	(0.048)		(0.045)	(0.044)	(0.047)	(0.053)	(0.139)
Number of male HH members	0.045		0.022	0.024	0.022	0.048	0.237*
	(0.053)		(0.051)	(0.051)	(0.058)	(0.061)	(0.132)
Any household shock last 12 months	−0.102		−0.077	−0.040	−0.027	−0.154	0.168
	(0.345)		(0.337)	(0.339)	(0.375)	(0.364)	(0.683)
=1 if Muslim	0.112			0.202	0.364	0.335	0.158
	(0.236)			(0.236)	(0.309)	(0.360)	(0.298)
=1 if unkown religion	0.925			0.822	0.926	−1.813	0.980
	(1.802)			(1.994)	(1.902)	(1.326)	(1.524)
=1 if forward caste	−0.312			−0.064	−0.385	−0.300	−0.776
	(0.309)			(0.306)	(0.319)	(0.347)	(0.581)

Continued on next page

Table 11: (Continued)

Sample: Children aged 0 to 5.							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
=1 if minority backward caste	−0.010			0.155	0.063	0.390	0.071
	(0.450)			(0.425)	(0.440)	(0.461)	(0.363)
=1 if scheduled caste or tribe	−0.317			−0.122	−0.356	−0.181	−0.518
	(0.246)			(0.256)	(0.265)	(0.281)	(0.401)
=1 if unkown Caste	0.391			0.502	0.457	0.049	1.532*
	(0.424)			(0.378)	(0.439)	(0.286)	(0.931)
Water and Garbage disposal Index	−0.445**						
	(0.182)						
Village Scale and Location Index					−0.095		
					(0.137)		
General Prices Index						−0.342	
						(0.481)	
Throw kitchen garbage away in waste baskets/trucks pick it up?							−0.752
							(0.684)
Does the village community get water for cooking and drinking from hand pump?							−0.080
							(0.267)
Does the village get water for cooking/drinking from household service connectio							−0.532
							(0.512)
Panel B: First Stage							
F-Stat	12.90	17.14	13.73	15.86	14.90	1.22	12.89
Obs	863	963	914	914	812	819	331
Clust	40	43	43	43	37	39	37
R2 Adj	0.10	0.09	0.07	0.08	0.10	0.04	0.10

SE clustered at Village level in parenthesis. Significance: * 10%, ** 5%, *** 1%.

A.3 Gender differences on nutrition

We can compare the nutritional inputs for children above the age of two years by gender. For these set of children we have information on the types of foods they consumed.[19] For this input we find no significant differences except for food items rich in proteins consumed by the child as part of its meal the day previous to the survey. This is an important food category within the context of this study. Puentes et al. [2014] for example show that "in contexts with substantial child malnourishment increases in protein-rich food intake in the first two years of life can have important effects on growth." However, our data suggests that it is girls that consumed a higher amount of proteins than boys, as shown in Table 12. This would suggest that differential nutrition investment is not at play in our setting. [20]

[19]This is the same information we used to construct the food diversity index earlier. The available data does not provide information on quantities consumed, only a yes/no indicator on consumption of different food items.

[20]We need to recall here that data on nutrition consumed by children above the age of two years is missing systematically, discussed earlier. The suggestive statements we make here need to be interpreted in this light and might not be valid for the representative household in the community.

Table 12: Nutritional inputs by gender

	Whole Sample	N	Female	Male	P-Value
Starchy staples	96.7	1262	96.8	96.6	0.429
Legumes	38.9	1229	37.9	39.9	0.349
Dairy (excluding breast milk)	87.0	1244	86.0	88.0	0.463
Meat, fish, egg	33.9	1237	37.0	30.9	0.014
Viamin A rich fruit or vegetables	44.5	1252	43.6	45.3	0.353
Other frutis or vegetables	42.5	1239	41.6	43.4	0.250
Foods made with oil, fats or butter	84.6	1246	84.7	84.5	0.515
Dietary diversity score of 0 to 2	8.0	1194	7.3	8.6	0.611
Dietary diversity score of 3 to 4	51.1	1194	52.9	49.3	0.131
Dietary diversity score of 5 to 7	41.0	1194	39.8	42.1	0.222

P-val corresponds to a t-test of difference of means between males and females clustering at village level. Round 2 data.