

Differences in Child Health Across Rural, Urban, and Slum Areas: Evidence From India

Claus C. Pörtner^{1,2} and Yu-hsuan Su³

RUNNING HEAD: Child Health Across Rural, Urban, and Slum Areas

Claus C. Pörtner (corresponding author)

e-mail: cportner@seattleu.edu

Yu-hsuan Su

e-mail: yuhstu@nccu.edu.tw

¹Department of Economics, Albers School of Business and Economics, Seattle University, 901 12th Avenue, Seattle, WA 98122, USA

²Center for Studies in Demography and Ecology University of Washington, Seattle, WA 98195, USA

³Graduate Institute of Development Studies, National Chengchi University, Taipei City 11605, Taiwan

Abstract The developing world is rapidly urbanizing, but an understanding of how child health differs across urban and rural areas is lacking. We examine the association between area of residence and child health in India, focusing on composition and selection effects. Simple height-for-age averages show that rural Indian children have the poorest health and urban children have the best, with slum children in between. With wealth or observed health environment held constant, the urban height-for-age advantage disappears, and slum children fare significantly worse than their rural counterparts. Hence, differences in composition across areas mask a substantial negative association between living in slums and height-for-age. This association is more negative for girls than boys. Furthermore, a large number of girls are “missing” in slums; we argue that this implies that the negative association between living in slums and health is even stronger than our estimate. The missing girls also help explain why slum girls appear to have a substantially lower mortality

than rural girls, whereas slum boys have a higher mortality risk than rural boys. We estimate that slum conditions (such as overcrowding and open sewers), which the survey does not adequately capture, are associated with 20 % to 37 % of slum children's stunting risk.

Keywords Child health; Slum; Urban; Rural; Sex selection

Introduction

Urban areas have a substantially lower percentage of stunted or underweight children than rural areas, but the absolute number of undernourished children has increased faster in urban than rural areas over the last decades (Fotso 2007; Haddad et al. 1999; Paciorek et al. 2013; Smith et al. 2005; van de Poel et al. 2007). As the developing world's urban population increases from 3.9 billion in 2011 to a projected 6.3 billion in 2050 (United Nations 2015:12), it is important that we understand how child health differs across rural and urban areas in order to design effective policies.

Why might child health differ across areas? One potential reason is the makeup of the population: on average, parents in urban areas are wealthier and better educated (Fields 1980), and such parents have healthier children (Strauss and Thomas 1995). If population composition explains the urban-rural child health differences, controlling for key determinants of child health should eliminate the urban advantage. However, the literature provides few conclusive results.

Controlling for composition reduces the urban advantage, but in many cases a statistically significant difference remains (Bocquier et al. 2011; Kennedy et al. 2006; van de Poel et al. 2009). For example, in prior research using Demographic and Health Surveys (DHS) data from 47 developing countries, differences in stunting and under-5 mortality risk between urban and rural areas remained statistically significant in 16 countries for stunting and 11 countries for mortality, even after a broad set of explanatory variables were controlled for (van de Poel et al. 2007; see also Dye 2008; Fotso 2006, 2007; Timæ us and Lush 1995; van de Poel et al. 2009).

Unequal distribution of wealth within urban areas further complicates the picture. In some cases, living in urban areas correlates with better child health for both wealthy and poor families, with the effect being larger the more well off the family (Dye 2008; Fotso 2006; Timæ us and Lush 1995). Some urban poor, however, live in environments and have health outcomes that are little better than those of the rural poor (Menon et al. 2000; Montgomery 2009). Furthermore, it is noteworthy that neighborhoods of relatively poor urban households are more heterogeneous than many believe (Montgomery and Hewett 2005). Finally, in some cases, the urban poor experience statistically significantly *higher* mortality than their rural counterparts when wealth and sociodemographic factors are controlled for (van de Poel et al. 2009).

A possible explanation for these inconclusive findings may be the failure to differentiate between slums and regular urban areas, given that descriptive studies have indicated that people in slums are less healthy compared with those in nonslum urban areas (Basta 1977; Ezeh et al. 2017; James et al. 1988; Mullick and Goodman 2005). Slums often serve as the first stop for people moving to cities in search of new opportunities; as the overall urban population grows, more and more people end up living in slums. Currently, approximately 863 million—or 33 %—of the urban population in developing countries live in slums (UN-Habitat 2013:151). However, research on slums and child health is scant, most likely because of a lack of data. DHS data, for example, include slum information for only three countries: Bangladesh, Egypt, and India. Most surveys exclude slum areas because they are often illegal settlements; when slum areas are included, the sample sizes are often too small to allow slum-specific estimates (Fotso 2007; Marx et al. 2013).

One way around the lack of household data is to examine the relationship between health outcomes and urban slum prevalence at the macro level. Country-level data from 80 developing countries show that a higher percentage of the population living in slums is associated with higher infant and child mortality (Jorgenson and Rice 2010, 2012; Jorgenson et al. 2012; Rice and Rice

2009). Another approach is to create slum indicators based on neighborhood characteristics from micro-level data. Using this approach on DHS data from 18 African countries, child mortality rates in slum areas are significantly higher than in nonslum urban areas, although in most cases, they are still lower than in rural areas (Günther and Harttgen 2012). Similarly, results using DHS data from 73 low- and middle-income countries show that when not controlling for determinants other than residence, slum children face higher health risks than urban children but lower risks than rural children (Fink et al. 2014). Controlling for maternal education, wealth, and health facilities access, health outcomes for slum children in towns with fewer than 1 million residents are not statistically different from those of rural children with comparable characteristics. What is more remarkable, however, is that slum children in cities with more than 1 million residents retain their health advantage over rural children, even when these controls are included.

These inconclusive results on the relationship between child health and area of residence provide the primary motivation for this article. We seek to discover the extent to which composition and selection effects explain differences in child health across rural, urban, and slum areas. Our child health outcomes are height-for-age, weight-for-height, and mortality risk: these outcomes capture different aspects of the same underlying, but unobserved, health-production process. We focus on height-for-age as our main child health indicator because it better measures children's long-term health and nutritional status and because higher height-for-age as a child is associated with more schooling and better labor market outcomes as an adult (Alderman et al. 2009; Maluccio et al. 2009; Thomas and Strauss 1997).

We take a different approach from the literature and focus on one country—India—for three reasons.⁴ First, India has the world's second-largest population and is home to substantial and

⁴ Other studies have examined rural-urban differences in specific countries but without allowing for the potential differences between slum and nonslum urban areas. For rural-urban mortality

rapidly growing slum areas. India's slum-dwelling population increased from 27.9 million in 1981 to 65.49 million in 2011 (India Office of the Registrar General and Census Commissioner 2013). India's largest city, Mumbai, has more than 6 million slum residents—of the city's total of 12 million people—even though slums occupy only approximately 9 % of the city's land. In addition, the number of slum dwellings in Mumbai has grown 40 % since 1995 (Murthy 2012).

Second, our data, the 2005–2006 National Family Health Survey, explicitly surveyed slum areas in addition to rural and urban areas. Having direct information on whether a respondent lives in a slum is important because constructing slum indicators based on household data, as done in the prior literature, is likely to miss important aspects. For example, population density and area conditions, such as open sewers, are defining characteristics of slums. However, most data sets do not provide information on either characteristic, making it difficult to successfully distinguish slum and nonslum areas using standard household data.⁵ The direct information on slums and the large sample size across all three types of areas—with slums oversampled to ensure a sufficient sample size—allow us to better identify potential differences across areas than the prior literature.

Finally, and most importantly, focusing on a specific country allows us to analyze potential selection effects in detail. We are particularly interested in selection effects that might arise from two sources: mortality and son preference.⁶

differences in India and Brazil, see Sastry (1997, 2004), Pradhan and Arokiasamy (2010), and Saikia et al. (2013). Studies that focus solely on slum areas include Subbaraman et al. (2012).

⁵ Günther and Harttgen (2012) and Fink et al. (2014) used the number of people per room as a proxy when defining slums, which may be more likely to pick up poverty than whether a household lives in a slum.

⁶ Other potential selection effects, such as migration, cannot be addressed because of data limitations.

Mortality selection is the potential for mortality to bias estimates of the association between area of residence and our height-for-age and weight-for-height measures of child health. We have information only on height and weight for children who survive to the survey date, but who survives is not random. Imagine a situation in which the distributions of underlying health for slum children and similar rural children are the same, but low-health rural children have a higher likelihood of dying than low-health slum children. In that case, rural children will *appear*, on average, to be healthier in terms of weight and height than slum children. A straightforward way to examine whether mortality selection is important is to compare our mortality results with our height and weight results: substantial differences suggest mortality selection.⁷

Mortality selection is a generic problem when estimating determinants of child health. A potentially bigger issue is selection from son preference. India's long history of strong son preference—especially in the northern states and among Hindus and Sikhs—manifests in higher mortality for girls than boys and prenatal sex selection.⁸

Two main concerns arise from the presence of son preference. First, selective recall often accompanies the differential mortality: deceased girls are less likely to be recorded than deceased boys when enumerators ask about fertility history (see Pörtner 2016: online appendix). The underreporting of female deaths leads to biased mortality estimates. Furthermore, because the deceased girls must have had worse health, on average, than those who survived, their deaths make

⁷ This approach does assume that the underlying health process is the same for all three outcomes and that the models are correctly specified. Pitt (1997) discussed estimating determinants of child health when there is potentially selection in fertility and mortality.

⁸ For an early discussion of son preference in India, see Sen (1990). See Pörtner (2016) for references on these different outcomes and an analysis of the relationships among fertility, birth spacing, and the use of sex selection.

the remaining population appear healthier. Hence, selective recall simultaneously worsens the mortality selection problem and makes it more difficult to establish whether it occurs, using comparisons of mortality and anthropometric results. If mortality and selective recall vary across areas, the estimates of the relationship between area of residence and child health are biased.

Second, any use of prenatal sex selection may also bias estimates of the relationship between area of residence and child health. Suppose we hold all observable characteristics constant and assume that a distribution of son preference exists across families and that the use of prenatal sex selection correlates positively with son preference after introduction of the technology. Then, all that is required for a bias is that girls not born because of prenatal sex selection would have suffered worse health and would have had higher likelihood of dying than girls born to families with the *same characteristics* but less strong son preference.¹⁰ Because the cost of raising children is higher in slum and nonslum urban areas than in rural areas, and because prenatal sex selection is therefore more prevalent (Pörtner 2016), estimated differences in child health and mortality across areas may suffer from bias, even when other characteristics are held constant.

Because son preference selection mainly affects girls, we estimate the relationships between area of residence and health outcomes both for boys and girls combined and separately for each sex.¹¹ This helps us understand whether the association between explanatory variables and health differ by sex and provides an indication of the extent of son preference selection.

Simple averages from National Family Health Survey 2005–2006 (NFHS-3) show that urban children fare better than slum children for all three health measures and that slum children

¹⁰ Empirical evidence exists for this mechanism in Taiwan, where access to sex-selective abortion reduced relative neonatal female mortality rates for higher-parity births (Lin et al. 2014).

¹¹ The absence of sisters and/or the expense of sex selection might affect the resources available to boys and therefore their health; even then, boys would be much less affected than girls.

fare better than rural children. However, after we control for wealth and health environment, no substantial difference in average height exists between urban (nonslum) and rural children, although slum children are significantly shorter than rural children. Hence, the composition of slum residents effectively hides the substantial negative association between living in slums and child health.

Controlling for wealth, health environment, and other observable characteristics, we find important differences between boys and girls. For example, the height difference between slum and rural children is larger for girls. Boys in slums do not have a higher probability of survival than rural boys, whereas slum girls appear to have a survival advantage over rural girls. We argue that the large number of missing girls in slums and urban areas indicates a substantial son preference selection, which makes mortality a poor measure for comparing health environments and biases the estimated negative coefficient on height for slums toward zero. We estimate that slum conditions—which we cannot adequately capture with currently available data—are associated with 20 % to 37 % of slum children’s risk of stunting.

Data and Estimation Strategy

We use data from the 2005–2006 National Family Health Survey (NFHS-3). NFHS-3 is the third in a series of national surveys, with earlier surveys in 1992–1993 (NFHS-1) and 1998–1999 (NFHS-2). We use NFHS-3 exclusively because the NFHS-1 did not include information on slums, and the NFHS-2 collected data from slum residents in Mumbai but not any other cities. The survey is described in detail in International Institute for Population Sciences (IIPS) and Macro International (2007: chapter 1).

In eight cities—Chennai, Delhi, Hyderabad, Indore, Kolkata, Meerut, Mumbai, and Nagpur—the NFHS-3 surveyed both urban nonslum areas and urban slum areas. NFHS-3 used two methods to identify areas as slum. The first method is the 2001 census classification of the area.

The census divided slums into three categories: “(i) all specified areas in a town or city notified as ‘Slum’ by State/Local Government and UT Administration under any Act including a ‘Slum Act’; (ii) all areas recognized as ‘Slum’ by State/Local Government and UT¹³ Administration, Housing and Slum Boards, which may have not been formally notified as slum under any act; and, (iii) a compact area of at least 300 population or about 60–70 households of poorly built congested tenements, in unhygienic environment usually with inadequate infrastructure and lacking in proper sanitary and drinking water facilities” (Gupta et al. 2009:10). The third category consists mainly of what is known as “nonnotified slums.”

The second method of slum identification is a local field supervisor assessment of whether a primary sampling unit (PSU) is in a slum.¹⁵ The definition that supervisors are asked to apply is equivalent to the third category used by the 2001 census and is meant to capture the less-established slums (see also Gupta et al. 2009:15). The assessment was collected for each surveyed PSU in the eight cities.

We consider a PSU a slum if it is identified as such by the 2001 census, the field supervisor, or both. We examine the sensitivity of our results to the choice of slum definition later in the paper.

To make the rural, slum, and urban samples more comparable, we restrict our sample to the seven states that have slum samples: Delhi, Uttar Pradesh, West Bengal, Madhya Pradesh, Maharashtra, Andhra Pradesh, and Tamil Nadu. Finally, we restrict the sample to Hindus and Muslims because of the very small number of surveyed slum children who are not Hindu or Muslim.

¹³ A union territory (UT) is an administrative unit in India, governed directly by the central government.

¹⁵ In urban areas, PSUs follow the 2001 census enumeration blocks, which contain 150–200 households. See IIPS and Macro International (2007:appendix C) for the selection process.

Estimation Strategy

To understand the role of composition effects, we begin by showing descriptive statistics for child health outcomes by area of residence. We then estimate a series of regressions with an expanding set of covariates. Our main indicator for child health is height-for-age z scores. A child with a z score of 0 is exactly the mean height of the comparison population for that age, and children with negative z scores are shorter. We also report the results for weight-for-height z scores.¹⁶

Our first model shows the association between area and child health when we control for child gender and age:

$$H_{ijk} = \alpha + \mathbf{A}_{jk}\beta_1 + \mathbf{C}_{ijk}\beta_2 + \epsilon_{ijk}, \quad (1)$$

where H_{ijk} is the health status of child i in household j in state k . \mathbf{A}_{jk} captures the area of residence of the household, divided into three exclusive areas: rural, urban nonslum, and urban slum. \mathbf{C}_{ijk} is a vector of personal characteristics of that child, including gender and age in months.

The second specification incorporates parental characteristics:

$$H_{ijk} = \alpha + \mathbf{A}_{jk}\beta_1 + \mathbf{C}_{ijk}\beta_2 + \mathbf{P}_{jk}\beta_3 + \mu_k + \epsilon_{ijk}, \quad (2)$$

where \mathbf{P}_{jk} is a vector of parental characteristics that includes mother's and father's levels of education, mother's height, the household head's religion, and the household head's caste.¹⁷ In addition, this specification includes fixed effects, μ_k , for state and survey month. We include survey month fixed effects to capture potential seasonal variation in child health, which is especially a concern in rural areas (IIPS and Macro International 2007). With the combination of state and

¹⁶ We do not use the information on diarrhea, cough, and fever because of the noisiness of these self-reported variables.

¹⁷ Following IIPS and Macro International (2007), the estimations use a set of dummy variables to capture parental education: 1–4, 5–7, 8–9, 10–11, and 12 or more years. Father's height is not included because the information is missing for more than one-half of the children in our sample.

month fixed effects, we capture the average child health for an area rather than simply the health status observed at the time of the survey.¹⁸

To examine the role of household wealth and observed local health environment, we then estimate the following:

$$H_{ijk} = \alpha + \mathbf{A}_{jk}\beta_1 + \mathbf{C}_{ijk}\beta_2 + \mathbf{P}_{jk}\beta_3 + \mathbf{W}_{jk}\beta_4 + \mathbf{R}_{jk}\beta_5 + \mu_k + \epsilon_{ijk}, \quad (3)$$

where \mathbf{W}_{jk} is a vector of the household's wealth status, and \mathbf{R}_{jk} is a vector of area characteristics for household jk . We include, in turn, household wealth, area wealth distribution, and area health environment, followed by a model that includes all three variables.

Each area characteristic variable is calculated using the “minus i ” method: $R_{jk} = \frac{1}{n-1} \hat{\alpha} R_{-jk}$,

where $\hat{\alpha} R_{-jk}$ indicates that the sum is over all other households in the PSU except for jk . The advantage is that area characteristics by construction are no longer correlated with the unobserved characteristics of the individual household (Aizer 2010). We capture household wealth using the NFHS-constructed wealth index, described later.

After examining the association between area of residence and height-for-age and weight-for-height, we turn to mortality. We examine mortality for two reasons. First, mortality is of interest in its own right as a health outcome. Second, child mortality results provide an indication of whether our height-for-age and weight-for-height results have any mortality selection problems. We estimate the association between the same sets of individual, household, and area characteristics and child mortality using the Cox proportional hazard model.

For a child i aged t months, the instantaneous hazard rate of death, conditional on still being

¹⁸ This does assume that the seasonal pattern is similar across states, but the loss of degrees of freedom if we were to interact month of survey with state would be large, and we would fail to capture the seasonal variation in health because no state survey covered the entire year.

alive at age t , is

$$\lambda(t|\mathbf{X}_{ijkt}) = \lambda_0(t)\exp(\mathbf{X}_{ijkt}\beta). \quad (4)$$

The baseline hazard, $\lambda_0(t)$, is a nonparametric time-varying function, and \mathbf{X}_{ijkt} is a vector that combines the explanatory variables in the previous specifications. For each child, the outcome is age—measured in months at the time of survey if still alive—or age of death if not alive at the time of the survey. An observation for a child alive at the time of the survey and less than 5 years of age is considered censored.

For all models and outcomes, we present results for three samples: all children combined, boys only, and girls only. All regressions employ survey weights to account for oversampling of slum areas. Furthermore, we use robust standard errors clustered at the PSU level for all regressions to allow for potential intragroup correlation of errors. We cluster at the PSU level because that is the highest level of aggregation for which we have variables of interest (Moulton 1990). All regressions are done in Stata 12.1 using the “cluster” option, which also implies robust estimation of the standard errors; we run Cox proportional hazard models using *stcox* and all other regressions using *regress*.

Variables and Descriptive Statistics

Table 1 presents descriptive statistics by area of residence: rural, slums, and urban nonslums. We limit the sample to children younger than age 5 because anthropometric information is not available for older children. Consistent with previous research, the overall health status of children in the sample is poor. The average height-for-age z score is -1.78 . Children in rural areas fare the worst, with an average height-for-age z score of -1.99 ; slum children have an average height-for-age z score of -1.59 ; and urban children are the healthiest, with a height-for-age z score of -1.50 . The differences across the three areas are all statistically significant at the 5 % level; the t statistics are 11.06 for rural–slum, 15.38 for rural–urban, and 2.36 for slum–urban. Using a threshold of

height-for-age z score of -2 , we find that more than one-half of the rural children are stunted, whereas approximately 40 % fall in this category for slums and urban areas. The difference in percentage stunted between urban and slum is not statistically significant. Hence, in line with prior research—and despite the common view of slums as detrimental to health—slum children fare surprisingly well according to the simple averages.

[place Table 1 about here]

Weight-for-height follows a similar pattern as height-for-age but with less distinct differences. The differences between rural and both slum and urban are statistically significant, but the difference between slum and urban is not; the t statistics are 7.82 for rural–slum, 9.84 for rural–urban, and 0.73 for slum–urban. Both the height-for-age z score and the weight-for-height z score are close to normally distributed and do not appear to be substantially skewed; see Online Resource 1 for histograms of outcomes by area and sex.

For mortality analyses, we expand the sample to include 1,118 children who died before their fifth birthday, making the sample 16,179 children born in the five years prior to the survey. As with the two other health measures, children in rural areas fare the worst, with a mortality rate of 8.3%; the mortality rate is 5.5 % in slums and 5.1 % in urban nonslum areas. Despite the relatively low mortality, the oversampling of slum populations helps ensure that we should have sufficient power; in slums, 167 of 3,138 children died, of which 72 were female. For comparison, 239 of 4,726 children in nonslum urban areas died, of which 98 were female. Mortality risk follows the same overall pattern in the three areas, with the majority of mortality concentrated within the first months of life, and almost no deaths after the first two years of life; see Online Resource 1 for nonparametric Kaplan-Meier survival curves using survey weights.

The natural sex ratio at birth in India is approximately 105 boys per 100 girls (Pörtner 2016). Hence, in the absence of differential mortality, sex-selective abortions, and selective recall

of deceased children, we should expect 48.8 % of the sample to be girls. The percentage of girls in rural areas is at the expected number. In urban areas, 48 % of the sample are girls, compared with only 46 % in slum areas. This provides a first indication that son preference selection may affect estimates of the association between area and child health.

The average levels of education for both mothers and fathers in urban non-slums and slum areas are substantially higher than in rural areas. Average educational levels between slum areas and urban non-slum areas differ by less than a year for both mothers and fathers. Corresponding to the height differences between children, mothers are, on average, tallest in urban areas, followed by slum areas, and finally by rural areas.

The wealth index in NFHS-3 is a composite measure of household living standard, based on principle components analysis of 33 assets and household characteristics. We use the wealth quintiles rather than the underlying index itself.²⁵ Not surprisingly, rural areas are the poorest, with 60 % of the children belonging to households in the bottom two wealth categories. Urban areas have the highest proportion in the top category (Category 5), with 47 % of children in that category, but slums are not far behind, with 38 % in the top category. Furthermore, 78 % of slum children belong to the top two wealth groups in slums, compared with 74 % in urban areas.

The bottom portion of Table 1 shows area wealth distribution and area health environment. All area characteristics are calculated as the average of households in the PSU, excluding the household itself, as described in the Estimation Strategy section. Area wealth distribution is captured by the percentage of households in each of the five wealth categories. As expected, given the distribution of wealth discussed earlier, slums and urban areas are relatively similar in terms of area wealth distribution, while households in rural areas generally have less wealth.

Area health environment includes characteristics that are thought to broadly reflect the

²⁵ See <http://dhsprogram.com/topics/wealth-index/> for an in-depth discussion of the wealth index.

healthiness of the living conditions of the area. These characteristics include water access (captured by the average time to fetch water and the type of drinking water source), access to improved cooking fuel (electricity, natural gas, biogas, and kerosene), sharing a toilet with 10 or more households, access to improved toilet facilities, and the average number of people per room.

The time it takes to fetch water—approximately six minutes—is close to identical across urban and slum areas, which is approximately one-half the time it takes in rural areas. Using the NFHS-3 report’s definition of access to improved sources of drinking water, approximately 96 % of households in urban and slum areas have access to an improved source of drinking water, with rural areas only slightly behind at 87 %.²⁶

Smoke from solid cooking fuels (coal/lignite, charcoal, wood, straw, shrubs, grass, agricultural crop waste, and dung cakes) is a serious health hazard, and we therefore include whether the household has access to improved cooking fuel (IIPS and Macro International 2007). The proportion of households that use improved cooking fuels is higher in slums (78 %) than in urban areas (69 %). Rural areas are far behind, with only 7 % using an improved cooking fuel.

At 19 %, slums have the highest percentage of households sharing toilets with 10 or more other households, probably because most slum dwellers rely on public toilets in the community. In urban areas, 6 % of households share with 10 or more households, while less than 1 % do so in rural areas. Approximately three-quarters of households in slums and urban areas have access to

²⁶ In addition to water piped into the dwelling, yard, or plot, an improved drinking water source includes water available from a public tap or standpipe, a tube well or borehole, a protected dug well, a protected spring, rainwater, and bottled water (IIPS and Macro International 2007). We also tried splitting into four main safe water sources, but none were statistically significantly different from unsafe/unimproved water sources. All showed coefficients close to 0, and the changes in the association between area dummy variable and height-for-age z score were minimal.

improved toilets, but only 17 % in rural areas have similar access.

Finally, slums and rural areas have essentially the same number of people per room at 3.7. Urban households have an average of 3.3 people per room.

Two important points arise from these descriptive statistics. First, normal measures of standard of living and area of residence are not necessarily closely correlated. For example, more children fall in the top two wealth categories in slum areas than in urban areas, and educational levels are relatively high. This finding is in line with prior research showing that slum residency is not equal to poverty and vice versa (Bhan and Jana 2013; Montgomery 2009; Montgomery and Hewett 2005). Second, even though the general perception of slums is one of squalor and poor living conditions, the descriptive statistics appear to paint a different picture. Differences in many household characteristics across areas are relatively small, and for some characteristics, slums even appear to fare best.

Results

The simple averages show that slum children—although clearly worse off than children in urban areas—do not lag far behind in terms of health and certainly are in better health than rural children. At question is the extent to which these simple averages provide an adequate description of the association between child health and area of residency. We begin by examining how composition effects influence child height-for-age. We then turn to mortality and son preference selection. Finally, we examine the robustness of our results.

Composition Effects and Height-for-Age

Table 2 presents the results for the aforementioned specifications for child height-for-age z scores.²⁸ Columns 1–7 show the results for different sets of control variables, beginning with the specification that includes only child age and sex and ending with the specification that includes all

²⁸ Online Resource 1 shows results by religion and caste affiliation.

variables. Only the estimated differences across areas and by sex are presented here (full results are available upon request).

[place Table 2 about here]

The simplest specification, column 1, which includes only age dummy variables, shows that children in urban slums appear to be taller than rural children, with children in urban areas the tallest. Compared with rural children, slum children are, on average, 0.38 standard deviations taller, and urban children are 0.43 standard deviations taller. Both differences are statistically significant at the 1 % level. Controlling for parental education, mother's height, household head religion and caste, and state and survey month fixed effects in column 2 substantially reduces the urban health advantage, and the difference between children's health in rural and slum areas is no longer statistically significant. After we include household wealth, area wealth distribution, or area health environment as an additional explanatory variable (shown in columns 3–5), living in slums is associated with statistically significantly worse health than living in rural areas. Furthermore, the difference is large: in the full specification, shown in column 7, a slum child is 0.22 standard deviations shorter than a rural child, with all other observable factors held constant.

Restricting the sample to boys only, we find no significant difference between rural and urban or rural and slums in columns 3–5 and column 7, although slum areas are still substantially below rural areas. The estimates for girls show only a very small difference between child health in rural and urban areas for columns 3–7. Living in slums is, however, associated with substantially and statistically significantly worse health relative to rural areas. After we control for wealth status, wealth distribution, or area health environment, girls in slums are almost one-quarter of a standard deviation shorter than girls in rural areas.

As in previous research on child health, the results for height-for-age are substantially stronger than those for weight-for-height (see Online Resource 1 for the results for weight-for-

height z scores). The overall pattern of the weight-for-height results is, however, strikingly similar to the height-for-age results. The basic specification shows an advantage in child health for both slums and urban children over rural children, with the largest difference for urban children. After we control for variables such as household wealth, area wealth distribution, and area health environment, slum children have lower weight-for-height z scores than rural children. The differences among slum, urban, and rural areas, however, are not statistically significant in the specifications that include area characteristics.

Overall, the findings suggest that a composition effect is at least partly responsible for the simple averages showing relatively healthy children in both urban and slum areas. Controlling for either household wealth or area characteristics, a child living in a slum is significantly shorter than what we would expect for a child with the same observable characteristics in a rural area, although this association should not be taken as causal. The question is whether mortality selection affects these results.

The Role of Mortality Selection

Table 3 follows the same specifications to examine how child mortality differs by area, except that age of the child is incorporated directly into the baseline hazard. The coefficients presented are hazard ratios; a coefficient less than 1 indicates a lower risk of death compared with the reference group, whereas a coefficient greater than 1 indicates a higher risk than the reference group. For the pooled sample of boys and girls, the simplest specification implies a hazard that is more than 40 % lower for children in slums and urban areas compared with rural areas, and these estimates are statistically significantly different from 1.

[place Table 3 about here]

Slum and urban children have substantially better survival chances than rural children. This pattern does not change when we include additional variables, although the additional variables

reduce the differences in survival chances across areas. Urban children have 20 % to 25 % lower mortality hazard than rural children with the same characteristics, and this difference is statistically significant in all models. Slum children have a similar or even higher advantage, but the estimate is outside the normal significance interval for the full model in column 7. These mortality results do appear to complicate the story. Using the pooled sample and the full models, slum children do significantly worse than rural children in terms of height, but slum children also have lower mortality, albeit not significantly so. Hence, it is possible that mortality selection partly explains the poorer health outcomes in slum areas when we take composition effects into account.

Countering this interpretation are the large differences in results by sex. Boys in slums have *higher* mortality and substantially worse height outcomes compared with rural boys; in urban areas, boys have statistically insignificantly worse height outcomes but lower mortality risk than rural boys. Thus, we find no clear evidence of selective mortality driving the health results for boys.

Girls show a distinctly different pattern from boys. In all specifications, girls from both slums and urban areas are substantially *less* likely to die than girls from rural areas. In fact, girls from slums appear to have identical or lower mortality than girls from urban areas in all specifications. In the full model, column 7, the mortality hazard for slum girls is 70 % lower than the hazard for rural girls, compared with a mortality hazard for urban girls of 40 %. Hence, part of the reason why we observe poorer health outcomes in slums may be due to the much lower observed mortality among girls in slums relative to rural girls.

The Role of Son Preference Selection

The mortality selection explanation for girls comes with an important caveat: both urban and slum areas show substantial bias in observed sex ratios at birth. Assuming that no boys are missing because of sex-selective abortions or selective recall errors, we should observe 100 / 105 girls born per boy born (Pörtner 2016). Hence, with 1,612 recorded male births in slums, we should expect

$1,612 \times (100 / 105) = 1,535$ girls born in slums. We observed, however, only 1,374 girls born in slums. Similarly, with 2,354 boys born in urban areas, there should be 2,242 girls born, but only 2,122 are observed. For rural areas, there are 4,456 male births and 4,261 female births, meaning that 51.1 % of the children born are boys, which corresponds closely to the expected sex ratio.

To get an idea of how much son preference selection affects our estimate, we combine observed mortality with the number of female births missing. This tells us for how many children we would have observed anthropometric information with no selection because of mortality, selective recall, or sex-selective abortions. If we restrict to girls, the combination of missing and dead over observed births and predicted missing would be 9 % in rural areas, 10 % in urban areas, and 15 % in slums.³¹ That is, we lack health information for a much higher proportion of children in slum areas than in rural or urban areas.

Seen in this light, our puzzling mortality results—with a very low mortality risk for slum girls relative to rural girls, but a higher mortality risk for slum boys relative to rural boys—make more sense. Girls in slums who were not observed because of son preference selection would likely have had both higher morbidity and higher mortality than what we see for the observed girls. In other words, a possible reason why mortality appears so low for girls in slums is that those at highest risk of dying are simply never recorded or born. Thus, the slum results on height-for-age and weight-for-height are likely underestimates: child health in slums compared with rural areas would be even worse if there were no son preference selection.

³¹ There are 384 female deaths of 4,261 female births in rural areas. In urban areas, there are 120 missing girls, 98 female deaths, and 2,122 observed female births. To calculate the percentage, we add the number of missing girls to the number of observed births to yield $(98 + 120) / (2,122 + 120) = 0.10$. Finally, for slum areas there are 161 missing girls, 72 female deaths, and 1,374 observed female births, yielding $(72 + 161) / (1,374 + 161) = 0.15$.

How Sensitive are Results to Slum Definition?

Four potential issues with respect to our slum definition can be identified. First, the lack of one correct objective definition of a slum (Bhan 2013), may explain the substantial idiosyncrasies by city in how supervisors classified slums relative to the census definition (see Gupta et al. 2009:73, table 1.1). At one end, supervisors in Indore agreed with only 5 of the 30 areas classified as slums by the census, and supervisors classified no additional PSUs as slums. At the other end, supervisors in Delhi agreed with all but four of the census slum PSUs and classified only two of the census nonslum areas as slums.

We classify an area as slum if either the census or the team supervisor indicated it as such. To examine how the slum definition affects our results, we replicated the estimations using two alternative definitions of slum: the 2001 census definition only or the supervisor definition only (results shown in Online Resource 1). The results for height-for-age using either the census definition or the supervisor's assessment correspond closely to those in Table 2. For the census definition, the full model coefficient for slum using the pooled sample is -0.17 , just outside the 10 % significance level. Comparatively, the supervisor's definition leads a point estimate of -0.19 , which is statistically significant at the 10 % level. Both coefficients are lower than the -0.22 we find, but not statistically significantly so.

Second, the 2001 census—on which the NFHS-3 sampling frame is based—identified slum and nonslum areas two to three years prior to the census, and some areas may therefore have changed status in the almost 10 years from the creation of the census frame to the NFHS-3 survey. The biggest concern is that we fail to capture some newer slum areas in the eight cities, especially because these newer slums are likely worse (Bhan and Jana 2013; Fink et al. 2014; Montgomery 2009; Subbaraman et al. 2012). The small and statistically insignificant differences between the results using census or supervisor definitions of slums help ease this concern. Even if the census

definition misses areas that have emerged as slums more recently, these would likely be captured by the supervisor during the survey, given that supervisors were asked to classify areas as slums if they fit the nonnotified/nonrecognized definition of slums.³⁴ A related concern is that some of the areas originally classified as slums by the census have developed enough that they no longer qualify. For both concerns, the effect would be underestimation of the strength of the negative association between slums and child health relative to rural areas.³⁵

Third, the representativeness of the data is a potential issue because the slum sample covered only the selected cities—although these eight cities did account for nearly 30 % of India’s slum population in 2001 (Gupta et al. 2009). How much this matters depends on two factors: (1) whether slums in the nonselected urban areas are different from the slums for which we do have data, and (2) whether slums in nonselected urban areas are still included in NFHS-3 but are captured as regular urban areas. Based on the (imperfect) slum measures created from household data used in prior research, slums in smaller urban areas may be worse than slums in larger urban areas (Fink et al. 2014). If so, our results are lower-bound estimates of the negative association between slums and child health.

If many slum areas were surveyed in the nonselected urban areas, this would downwardly bias our estimated association between living in urban nonslum areas and child health, making urban nonslum areas seem unhealthier than they really are. Unfortunately, there is no direct way to establish the extent to which slums in other urban areas than the eight cities were surveyed. We can, however, split urban areas into the eight selected for the slum survey and those that were not—and

³⁴ The caveat to this argument is that supervisors might have been too stringent and therefore also failed to classify areas as slums.

³⁵ It would also make urban nonslum areas appear less healthy relative to rural areas because any newer, missed slums would be classified as urban areas.

therefore perhaps include some slum areas—and reestimate the models (results shown in Online Resource 1). The weight-for-height results are practically identical across selected and nonselected urban areas, and the slum results do not change. Of the possible explanations for these results, we consider it most likely that NFHS-3 covered very few slums in the nonselected urban areas, combined with little difference in the association between living in urban areas and child health across the different urban areas. Another possibility is that some slums were surveyed, but not recorded as slums in the nonselected urban areas, combined with the other parts of the nonselected urban areas being substantially healthier than the selected urban areas.

Finally, there clearly is the potential for variation across slums in how unhealthy they are. For example, survey data from Kaula Bandar, a nonnotified slum in Mumbai, show relatively worse health outcomes compared with NFHS-3 slum data from Mumbai, likely because of Kaula Bandar's nonnotified status (Subbaraman et al. 2012). We cannot address this important topic because we cannot reliably identify different types of slums in the data. Ultimately, however, these concerns point to the association between living in slums and health being even more negative than our estimates show.

Other Selection or Specification Issues?

Finally, in addition to selection from mortality and son preference, other selection issues or omitted variable biases may affect our results. The main candidate is selective migration. For example, if parents believe that slums are bad for child health, those parents who care the most about child health are the most likely to not live in slums, but their children would fare better under any circumstances. In that case, slums would seem worse than they really are because the parents remaining in slums care less about child health and therefore have worse outcomes.

All the limited information on migration in NFHS-3 allows us to do is identify two groups: migrants (those not born in the neighborhood in which they are surveyed) and nonmigrants (those

who have never moved). Despite the common perception that slums are mainly populated by a transient population, the percentage of mothers born in the neighborhood in which they are interviewed is higher in slums (27 %) than in urban non-slums (22 %) and rural areas (14 %).³⁸ The distribution is in line with Fry et al.'s (2002) argument that slums often are stable and homogeneous communities rather than chaotic agglomerations, although see the discussion of slums as poverty traps in Marx et al. (2013). The low number for the rural population is most likely the result of the Indian practice of exogamy, in which a woman marries into a household in another village and becomes part of her husband's household (Rosenzweig and Stark 1989). The possibility of selective migration points to the importance of collecting detailed information on migration behavior in future surveys so that researchers may better understand household migration decisions.

Finally, our results are conditional on including a set of covariates that best eliminate omitted variable bias and correctly specifying the regression models. We have expanded on the set of covariates used compared with prior research in this field, but other unobserved variables could be correlated with both our chosen covariates and the child health outcomes. The stability of the results for different combinations of covariates, however, is encouraging. We discuss the need for better survey information, especially for area characteristics, in the Conclusion.

³⁸ These numbers are based on the question asked of all women: "How long have you been living continuously in (NAME OF CURRENT PLACE OF RESIDENCE)?," where name of current place of residence is the village's name in rural areas and the neighborhood in urban areas. Hence, we can reasonably expect that a woman surveyed in a slum who responds that she was born in the same neighborhood would have spent her entire life in the slum. For more on this question, see the discussion on DHS user forum on India

(https://userforum.dhsprogram.com/index.php?t=tree&goto=11187&&t=tree&goto=11187&##msg_11187). Results by group are available upon request.

Discussion

Our finding of a substantial negative association between slums and height-for-age when controlling for household characteristics runs counter to some of the recent literature, especially Fink et al.'s (2014) cross-country analysis. There are minor differences in outcomes and estimation methods used in our study versus previous studies. We focus on height-for-age z scores, rather than the simple cutoff of stunting, and use an expanded set of explanatory variables. Neither difference, however, is likely to explain the differences in results.

Why then do the results differ? One important reason may be that NFHS-3 designed the sample frame to incorporate slums and provides a slum indicator. Previous studies had to create slum indicators based solely on information about households in the areas, which cannot capture area characteristics such as overcrowding and unhygienic local conditions.³⁹ Furthermore, other surveys may not even include slum areas if these areas are not explicitly targeted (Fotso 2007; Marx et al. 2013). If the areas designated as slums in the prior literature are not really slums but rather simply poorer urban areas, this may explain why prior studies failed to find a difference across areas.

Another reason for the differences in results is that focusing on one country and estimating results for boys and girls separately allow us to better examine how selection issues affect results. The very low mortality risk for slum girls and the large number of missing girls in slums point to a potential role for son preference selection.⁴⁰ If son preference selection differs by area, mortality

³⁹ Consider, for example, using the number of people per room as an indicator for crowding. The number of people per room fails to capture that dwellings in slums are located much closer together than in either urban or rural areas, and the average number of people per room varies little across areas.

⁴⁰ The main caveat is that we cannot rule out selective migration.

results contribute little to our understanding of how child health differs across areas. The existence of son preference suggests that our estimates are lower bounds and that without son preference selection, the slum estimates would have been even worse than what we find. This leaves the question of what explains the negative association between slums and child health. The slum dummy variable captures the average difference in child health between slums and rural areas, *conditional* on the observable characteristics in our regressions. We expect that the broader, unobserved health environment of slums explains most of this difference. Three factors are likely the most important components of this health environment: open sewers, overcrowding, and poor water quality. All three are either insufficiently captured by DHS data or not captured at all.

Water quality is a particularly interesting example because slums appear to have better access to improved water sources than either urban nonslum or rural areas. Our water access variables provide, however, only an imprecise measure of actual water quality because they do not consider the reliability of the supply (Satapathy 2014). Intermittent water supply reduces water quality because interruptions in supply allow contaminants (such as human excreta) to enter the pipes, with the contaminants then distributed across the system when the supply is restored. Even with identical supply interruptions across urban and slum areas, water quality would likely be worse in slums because of overcrowding, open defecation, and poor sewage systems. Low water quality affects child health through environmental enteric dysfunction, where contaminated water or other environmental factors change children's gut bacteria, leaving them more prone to malnutrition despite being fed what appears to be an adequate diet (Keusch et al. 2013). This mechanism may also explain much of the variation in rural height-for-age between India and Africa (Spears 2013). Furthermore, evidence suggests that exposure to open defecation is increasing in India (Spears, 2014).

In summary, we find a negative association between living in slums and child health

because the broader health environment, which we cannot adequately capture in our data, is responsible for the lower levels of health in slums when we control for parental and observable area characteristics. The combination of poor water quality, open sewers, and overcrowding is a likely candidate.⁴²

Conclusion

The primary aim of this article is to examine the association between child health and residence area type. Simple averages from the third round of the India NFHS show the worst child health in rural areas and the best in urban areas, with slums in between. This finding runs counter to the common belief that slums are very unhealthy, but is in line with prior cross-country findings. The simple height-for-age averages, however, do not consider composition and selection effects, which may obscure an area's true health effects. Our main finding is a strong negative association between living in slums and children's height-for-age after we control for wealth or area characteristics. The negative association between slum residence and height-for-age is larger for girls than boys. Furthermore, selection effects are important. Mortality appears to be low for girls in slums, but this finding obscures that many girls are "missing," possibly because of selective recall of deceased girls or outright sex selection. Thus, the negative association between living in slums and health would likely be substantially worse for girls—and therefore overall—if we were

⁴² See also Bhan and Jana (2013). This may also explain why the positive relationship between mother's education and child health found in rural areas diminishes, or even disappears, in slum and urban areas, as shown in Online Resource 1. One interpretation is that slums' broader health environment is so bad that more education does little to counter the negative effects. That a mother knows to wash her hands, boil water before use, and take a sick child to the doctor matters little for child health if the local playground is an open sewer, or if diseases spread quickly and easily due to overcrowding.

somehow able to capture health outcomes for girls who either died or who were never born because of sex selection. Working in the same direction is that slums in smaller urban areas and perhaps newer slums in the selected cities are absent from our data. The caveats to our results are that we cannot address selective migration and that the results are conditional on correctly specified models.

To provide an idea of how living in slums is associated with child health, Table 4 shows the predicted percentage stunted by area and the predicted percentage stunted if slum children had the same risk of stunting as children in urban areas or rural areas—but otherwise retained their other observed characteristics—using the pooled sample, girls only, and boys only.⁴³ We consider a child stunted if the predicted height-for-age z score is -2 or lower, and we base our calculation of the predicted number of stunted children by area on our results, combined with the 2011 census.

[place Table 4 about here.]

Almost 51 % of children are stunted overall. Slums' contribution to overall stunting might seem insignificant because India is still a predominately rural society, and a substantial level of stunting exists in rural areas. Using the population attributable fraction approach, we find that only

⁴³ These numbers should be taken as suggestive at best and not as causal estimates, and are conditional on correctly specifying the underlying model with the aforementioned caveats.

Furthermore, although we do employ the weights provided in NFHS-3, we use only the subset of states that have slum information in NFHS-3, and the composition of the population in NFHS-3 may vary from India as a whole. Our total predicted number of stunted does not match that of, for example, UNICEF (2013) because the census count for slums included children aged 5 and 6.

Scaling our estimate by $5 / 7$ to get an approximation for those under 5 leads to a total number of 59.8 million stunted children below age 5, which is slightly lower than the 61.7 million quoted in the UNICEF report.

0.6 % to 1.2 % of India's stunting is associated with living in slums.⁴⁴ This percentage, however, obscures the fact that the predicted number of stunted children decreases by one-half million if slum children had similar risk as urban areas, and by 1 million if the risk was equal to rural areas. Hence, when we focus on slum children and control for observable characteristics, the population attributable fraction shows that 20 % to 37% of slum children's stunting risk is associated with unobserved slum conditions.

An important implication of our results is that the health environment variables generated from standard household surveys are unable to fully capture the differences in health environment, and fall particularly short for slums. Differences in many household characteristics across areas are relatively small, and for some characteristics, slums even appear to do best. Thus, in addition to more surveys that explicitly target a representative sample of slums with a sufficient number of observations, we need better measures of area characteristics—first and foremost in DHS surveys because of their extensive use in the analysis of child health. When trying to understand what exactly makes slums unhealthy, we would benefit from better area measures that would allow us to consider factors such as overcrowding, access to health services, sewage system quality, and reliability and contamination of water supply as contributors to poor slum health.

Until we have better data on area characteristics, our best guess for what explains the substantial negative association between living in slums and child health is a combination of unreliable water supply, open sewers, and overcrowding. This combination results in low water quality, leading to environmental enteric dysfunction and poor health outcomes, even when other household characteristics suggest that the child should fare relatively well. Thus, policies that emphasize physical infrastructure—such as reliability of water supply—might prove more cost-

⁴⁴ See Levine (2007) on calculation and interpretation of the population attributable fraction.

effective than those focusing on changing household behavior and characteristics. The perennial problem, of course, is that the very nature of slums and the illegality of many dwellings make this difficult (Subbaraman et al. 2012). Furthermore, it is important to distinguish between poverty and slum targeting when designing policies (Bhan and Jana 2013). Both are important but are likely to lead to very different policies.

Our results also have broader implications for future research. Differences in the number of missing girls across areas are associated with mortality numbers that do not adequately reflect how health conditions differ and downwardly bias the estimated differences in health among areas. Understanding and addressing this selection effect when estimating mortality and health determinants is an important area for future research. This is especially the case for countries with strong son preferences, such as India and China.

The selection issues also provide a cautionary note on the use of cross-country data. Because of the large number of DHS data sets, with their ready availability and similar variable definitions across countries, researchers can now combine data from many countries for analysis. The combined data's large sample size means that we can address questions for which individual country-level samples may be too small. Cross-country data, however, also make adequately addressing country-specific factors—such as son preference—more difficult, potentially leading to biased results. One example is the prior finding that slums are not associated with worse child health, which we argue comes partly from this type of bias.

In conclusion, with slums associated with stunting of up to 1 million Indian children, and with the rapid increase in the developing world's urban population, understanding how child health differs across areas—and more generally, what determines child health in cities—is an undertaking with important policy implications, and one that will only become more important over time.

Acknowledgments We thank Seik Kim, Robert Plotnick, Judith Thornton, five anonymous referees, and participants at the Population Association of America annual meetings, Pacific Conference for Development Economics, Annual Conference of the European Society for Population Economics, DIAL Development Conference, and the Labor and Development Seminar at the University of Washington for their helpful comments. Partial support for this research came from a research grant from the Eunice Kennedy Shriver National Institute of Child Health and Human Development (R24 HD042828) to the Center for Studies in Demography & Ecology at the University of Washington, from the Office of Research and Development of National Chengchi University, and from the Ministry of Science and Technology of Taiwan Government (104-2914-I-004-009-A1).

References

- Aizer, A. (2010). The gender wage gap and domestic violence. *American Economic Review*, 100, 1847–1859.
- Alderman, H., Hoozevee, H., & Rossi, M. (2009). Preschool nutrition and subsequent schooling attainment: Longitudinal evidence from Tanzania. *Economic Development and Cultural Change*, 57, 239–260.
- Basta, S. S. (1977). Nutrition and health in low income urban areas of the third world. *Ecology of Food and Nutrition*, 6, 113–124.
- Bhan, G. (2013). Planned illegalities—Housing and the “failure” of planning in Delhi: 1947–2010. *Economic & Political Weekly*, 48(24), 58–70.
- Bhan, G., & Jana, A. (2013). Of slums or poverty. *Economic & Political Weekly*, 48(18), 13–16.
- Bocquier, P., Madise, N. J., & Zulu, E. M. (2011). Is there an urban advantage in child survival in sub-Saharan Africa? Evidence from 18 countries in the 1990s. *Demography*, 48, 531–558.
- Dye, C. (2008). Health and urban living. *Science*, 319, 766–769.

- Ezeh, A., Oyeboode, O., Satterthwaite, D., Chen, Y-F., Ndugwa, R., Sartori, J., . . . Lilford, R. J. (2017). The history, geography, and sociology of slums and the health problems of people who live in slums. *Lancet*, 389, 547–558.
- Fields, G. S. (1980). Education and Income Distribution in Developing Countries: A Review of the Literature. In T. King (Ed.). *Education and income: A background study for world development* (pp. 231-315). Washington, D.C.: The World Bank.
- Fink, G., Günther, I., & Hill, K. (2014). Slum residence and child health in developing countries. *Demography*, 51, 1175–1197.
- Fotso, J.-C. (2006). Child health inequities in developing countries: Differences across urban and rural areas. *International Journal for Equity in Health*, 5:9.
- Fotso, J.-C. (2007). Urban-rural differentials in child malnutrition: Trends and socioeconomic correlates in sub-Saharan Africa. *Health & Place*, 13(Special Issue), 205–223.
- Fry, S., Cousins, B., & Olivola, K. (2002, May). *Health of children living in urban slums in Asia and the Near East: Review of existing literature and data* (Environmental Health Project Activity Report 109). Washington, DC: U.S. Agency for International Development.
- Günther, I., & Harttgen, K. (2012). Deadly cities? Spatial inequalities in mortality in sub-Saharan Africa. *Population and Development Review*, 38, 469–486.
- Gupta, K., Arnold, F., & Lhungdim, H. (2009). *Health and living conditions in eight Indian cities*. National Family Health Survey (NFHS–3), India, 2005–06. Mumbai, India: International Institute for Population Sciences.
- Haddad, L., Ruel, M. T., & Garrett, J. L. (1999). Are urban poverty and undernutrition growing? Some newly assembled evidence. *World Development*, 27, 1891–1904.
- India Office of the Registrar General and Census Commissioner. (2013). *Primary census abstract for slum* (Technical Report). New Delhi, India: Office of the Registrar General and Census

Commissioner.

International Institute for Population Sciences (IIPS) and Macro International. (2007). *National Family Health Survey (NFHS-3), 2005-06: India, Vol. 1*. Mumbai, India: International Institute for Population Sciences.

James, W. P. T., Ferro-Luzzi, A., & Waterlow, J. C. (1988). Definition of chronic deficiency in adults—Report of a working party of the International Dietary Energy Consultative Group. *European Journal of Clinical Nutrition*, 42, 969–981.

Jorgenson, A. K., & Rice, J. (2010). Urban slum growth and human health: A panel study of infant and child mortality in less-developed countries, 1990–2005. *Journal of Poverty*, 14, 382–402.

Jorgenson, A. K., & Rice, J. (2012). Urban slums and children's health in less-developed countries. *Journal of World-Systems Research*, 18, 103–116.

Jorgenson, A. K., Rice, J., & Clark, B. (2012). Assessing the temporal and regional differences in the relationships between infant and child mortality and urban slum prevalence in less developed countries, 1990–2005. *Urban Studies*, 49, 3495–3512.

Kennedy, G., Nantel, G., Brouwer, I. D., & Kok, F. J. (2006). Does living in an urban environment confer advantages for childhood nutritional status? Analysis of disparities in nutritional status by wealth and residence in Angola, Central African Republic and Senegal. *Public Health Nutrition*, 9, 187–193.

Keusch, G. T., Rosenberg, I. H., Denno, D. M., Duggan, C., Guerrant, R. L., Lavery, J. V., . . . Brewer, T. (2013). Implications of acquired environmental enteric dysfunction for growth and stunting in infants and children living in low- and middle-income countries. *Food and Nutrition Bulletin*, 34, 357–364.

Levine, B. (2007). What does the population attributable fraction mean? *Preventing Chronic*

Disease, 4(1), 1–5.

- Lin, M.-J., Liu, J.-T., & Qian, N. (2014). More missing women, fewer dying girls: The impact of sex-selective abortion on sex at birth and relative female mortality in Taiwan. *Journal of the European Economic Association*, 12, 899–926.
- Maluccio, J. A., Hoddinott, J., Behrman, J. R., Martorell, R., Quisumbing, A. R., & Stein, A. D. (2009). The impact of improving nutrition during early childhood on education among Guatemalan adults. *Economic Journal*, 119, 734–763.
- Marx, B., Stoker, T., & Suri, T. (2013). The economics of slums in the developing world. *Journal of Economic Perspectives*, 27, 187–210.
- Menon, P., Ruel, M. T., & Morris, S. S. (2000). Socio-economic differentials in child stunting are consistently larger in urban than in rural areas. *Food and Nutrition Bulletin*, 21, 282–289.
- Montgomery, M. R. (2009). Urban poverty and health in developing countries. *Population Bulletin*, 64(2), 1–16.
- Montgomery, M. R., & Hewett, P. C. (2005). Urban poverty and health in developing countries: Household and neighborhood effects. *Demography*, 42, 397–425.
- Moulton, B. R. (1990). An illustration of a pitfall in estimating the effects of aggregate variables on micro units. *Review of Economics and Statistics*, 72, 334–338.
- Mullick, M. S. I., & Goodman, R. (2005). The prevalence of psychiatric disorders among 5–10 year olds in rural, urban and slum areas in Bangladesh. *Social Psychiatry and Psychiatric Epidemiology*, 40, 663–671.
- Murthy, S. L. (2012). Land security and the challenges of realizing the human right to water and sanitation in the slums of Mumbai, India. *Health and Human Rights*, 14(2), 61–73.
- Paciorek, C. J., Stevens, G. A., Finucane, M. M., & Ezzati, M. (2013). Children's height and weight in rural and urban populations in low-income and middle-income countries: A

- systematic analysis of population-representative data. *Lancet Global Health*, 1, e300–e309.
doi:10.1016/S2214-109X(13)70109-8
- Pitt, M. M. (1997). Estimating the determinants of child health when fertility and mortality are selective. *Journal of Human Resources*, 32, 129–158.
- Pörtner, C. C. (2016). *Sex-selective abortions, fertility, and birth spacing* (World Bank Policy Research Working Paper No. 7189). Washington, DC: The World Bank.
- Pradhan, J., & Arokiasamy, P. (2010). Socio-economic inequalities in child survival in India: A decomposition analysis. *Health Policy*, 98, 114–120.
- Rice, J., & Rice, J. S. (2009). The concentration of disadvantage and the rise of an urban penalty: Urban slum prevalence and the social production of health inequalities in the developing countries. *International Journal of Health Services*, 39, 749–770.
- Rosenzweig, M. R., & Stark, O. (1989). Consumption smoothing, migration, and marriage: Evidence from rural India. *Journal of Political Economy*, 97, 905–926.
- Saikia, N., Singh, A., Jasilionis, D., & Ram, F. (2013). Explaining the rural-urban gap in infant mortality in India. *Demographic Research*, 29(article 18), 473–506.
doi:10.4054/DemRes.2013.29.18
- Sastry, N. (1997). What explains rural-urban differentials in child mortality in Brazil? *Social Science & Medicine*, 44, 989–1002.
- Sastry, N. (2004). Urbanization, development and under-five mortality differentials by place of residence in São Paulo, Brazil, 1970–1991. *Demographic Research*, S2(article 14), 355–386.
doi:10.4054/DemRes.2004.S2.14
- Satapathy, B. K. (2014). Safe drinking water in slums: From water coverage to water quality. *Economic & Political Weekly*, 49(24), 50–55.
- Sen, A. (1990, December 20). More than 100 million women are missing. *New York Review of*

- Books*, 37. Retrieved from <http://www.nybooks.com/articles/1990/12/20/more-than-100-million-women-are-missing/>
- Smith, L. C., Ruel, M. T., & Ndiaye, A. (2005). Why is child malnutrition lower in urban than in rural areas? Evidence from 36 developing countries. *World Development*, 33, 1285–1305.
- Spears, D. (2013). *How much international variation in child height can sanitation explain?* (World Bank Policy Research Working Paper No. 6351). Washington, DC: The World Bank.
- Spears, D. (2014). *Increasing average exposure to open defecation in India, 2001–2011* (Working Paper 21). New Delhi, India: Research Institute for Compassionate Economics.
- Strauss, J., & Thomas, D. (1995). Human resources: Empirical modeling of household and family decisions. In Behrman, J., & Srinivasan, T. N. (Eds.). *Handbook of Development Economics* (pp. 1883–2023). Amsterdam: Elsevier Science.
- Subbaraman, R., O’Brien, J., Shitole, T., Shitole, S., Sawant, K., Bloom, D. E., & Patil-Deshmukh, A. (2012). Off the map: The health and social implications of being a non-notified slum in India. *Environment and Urbanization*, 24, 643–663.
- Thomas, D., & Strauss, J. (1997). Health and wages: Evidence on men and women in urban Brazil. *Journal of Econometrics*, 77, 159–185.
- Timæ us, I. M., & Lush, J. (1995). Intra-urban differentials in child health. *Health Transition Review*, 5, 163–190.
- UN-Habitat. (2013). *State of the world’s cities 2012/2013: Prosperity of cities*. New York, NY: Routledge.
- UNICEF. (2013). *Improving child nutrition: The achievable imperative for global progress*. New York, NY: UNICEF.
- United Nations. (2015). *World urbanization prospects: The 2014 revision*. New York, NY: United Nations, Department of Economic and Social Affairs,

Population Division.

van de Poel, E., O'Donnell, O., & van Doorslaer, E. (2007). Are urban children really healthier?

Evidence from 47 developing countries. *Social Science & Medicine*, 65, 1986–2003.

van de Poel, E., O'Donnell, O., & van Doorslaer, E. (2009). What explains the rural-urban gap in

infant mortality: Household or community characteristics? *Demography*, 46, 827–850

Table 1 Descriptive statistics

Variable	Rural	Slum	Urban	Total	<i>t</i> Tests ^a
Dependent Variables					
Height-for-age <i>z</i> score	-1.99 (1.67)	-1.59 (1.63)	-1.50 (1.69)	-1.78 (1.68)	ABC
Ratio with height-for-age <i>z</i> score < -2 ^b	0.51 (0.50)	0.40 (0.49)	0.39 (0.49)	0.46 (0.50)	AB
Weight-for-height <i>z</i> score	-1.04 (1.29)	-0.81 (1.35)	-0.79 (1.38)	-0.93 (1.33)	AB
Ratio with weight-for-height <i>z</i> score < -2 ^b	0.20 (0.40)	0.17 (0.38)	0.16 (0.37)	0.19 (0.39)	AB
Percentage who died ^c	8.28	5.32	5.06	6.82	N/A
Child Characteristics					
Girl	0.49	0.46	0.48	0.48	A
Age of child (months)	30.0 (17.0)	30.5 (16.8)	30.6 (16.7)	30.3 (16.9)	
Parental Education and Mother's Height					
Mother's education (years)	3.3 (4.3)	6.7 (5.0)	7.3 (5.6)	5.1 (5.2)	ABC
Father's education (years)	6.0 (5.0)	7.9 (4.8)	8.6 (5.4)	7.1 (5.2)	ABC
Mother's height (cm)	151.2 (5.6)	152.1 (5.7)	152.4 (5.8)	151.7 (5.7)	ABC
Household Religion and Caste					
Hindu	0.87	0.73	0.76	0.81	ABC
Muslim	0.13	0.27	0.24	0.19	ABC
Scheduled caste	0.23	0.20	0.17	0.21	ABC
Scheduled tribe	0.10	0.04	0.02	0.06	ABC
Other backward classes	0.44	0.34	0.40	0.41	ABC
Not a scheduled caste/tribe or backward class	0.23	0.42	0.40	0.31	AB
Household Wealth Status					
Wealth category 1 (poorest)	0.33	0.01	0.04	0.19	ABC
Wealth category 2	0.27	0.04	0.17	0.17	ABC
Wealth category 3	0.22	0.17	0.15	0.19	AB
Wealth category 4	0.14	0.40	0.27	0.22	ABC
Wealth category 5 (richest)	0.04	0.38	0.47	0.22	ABC
Area Wealth Distribution					
Ratio in wealth category 1 (poorest)	0.34	0.01	0.03	0.19	ABC
Ratio in wealth category 2	0.29	0.05	0.07	0.18	ABC
Ratio in wealth category 3	0.21	0.17	0.15	0.18	ABC
Ratio in wealth category 4	0.12	0.40	0.28	0.22	ABC
Ratio in wealth category 5 (richest)	0.04	0.38	0.47	0.22	ABC
Area Health Environment					
Average time to get water and return (minutes)	11.62	6.73	6.29	9.20	AB
Ratio with improved source of drinking water	0.87	0.96	0.96	0.91	AB
Ratio sharing toilet with 10 plus households	0.00	0.19	0.06	0.05	ABC
Ratio with improved toilet	0.17	0.74	0.75	0.44	ABC
Average number of people per room	3.66	3.68	3.33	3.57	BC
Ratio with improved cooking fuel	0.07	0.78	0.69	0.38	ABC
Number of Observations	7,993	2,822	4,246	15,061	

Notes: Standard deviations are shown in parentheses (not shown for categorical variables except stunting and wasting). All area variables are calculated as the mean for the household's PSU, excluding the household itself. There are, on average, 1.36 births per woman within the five years prior to the survey date, with a minimum of 1 and a maximum of 4.

^a *t* tests for differences between areas assuming unequal variance. A indicates a statistically significant difference at the 5 % level between rural and slum areas; B indicates a statistically significant difference at the 5 % level between rural and

urban areas; and C indicates a statistically significant difference at the 5 % level between slum and urban areas.

^b The ratio of stunted and wasted children is presented for information only and not used as dependent variables.

^c Mortality calculation is based on all births that occurred in the five years prior to the survey. The sample sizes are 8,717 in rural areas, 2,986 in slums, and 4,476 in urban non-slums, for a total of 16,179 births. The outcome used for Cox mortality regressions is length of life rather than a dummy variable for died.

Table 2 Covariates of child health: Height-for-age *z* score

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Child Age Dummy Variables ^a	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State and Survey Month Fixed Effects	No	Yes	Yes	Yes	Yes	Yes	Yes
Parental Education and Mother's Height ^b	No	Yes	Yes	Yes	Yes	No	Yes
Household Religion and Caste ^c	No	Yes	Yes	Yes	Yes	Yes	Yes
Household Wealth Status ^d	No	No	Yes	No	No	Yes	Yes
Area Wealth Distribution ^e	No	No	No	Yes	No	Yes	Yes
Area Health Environment ^f	No	No	No	No	Yes	Yes	Yes
All Children (<i>n</i> = 15,061)							
Slum	0.38** (0.06)	0.07 (0.06)	-0.12* (0.06)	-0.15 [†] (0.08)	-0.22* (0.11)	-0.26* (0.11)	-0.22* (0.11)
Urban	0.43** (0.06)	0.14** (0.04)	-0.03 (0.05)	-0.08 (0.07)	-0.08 (0.07)	-0.13 [†] (0.07)	-0.10 (0.07)
Girl	0.01 (0.03)	0.01 (0.03)	0.01 (0.03)	0.01 (0.03)	0.01 (0.03)	0.01 (0.03)	0.01 (0.03)
<i>R</i> ²	.10	.20	.21	.20	.20	.18	.21
Adjusted <i>R</i> ²	.10	.20	.21	.20	.20	.18	.21
Wald test against model		(1)	(2)	(2)	(2)	(7)	(2)
<i>F</i> statistics		50.03	23.69	7.88	3.28	33.55	7.44
Degrees of freedom		(15,1515)	(4,1515)	(4,1515)	(9,1515)	(11,1515)	(17,1515)
<i>p</i> value		.00	.00	.00	.00	.00	.00
Boys (<i>n</i> = 7,858)							
Slum	0.39** (0.08)	0.12 (0.08)	-0.09 (0.09)	-0.10 (0.11)	-0.19 (0.13)	-0.25* (0.13)	-0.19 (0.13)
Urban	0.41** (0.06)	0.13* (0.05)	-0.06 (0.06)	-0.10 (0.09)	-0.11 (0.09)	-0.17* (0.09)	-0.13 (0.09)
<i>R</i> ²	.09	.20	.21	.21	.21	.18	.21
Adjusted <i>R</i> ²	.09	.20	.21	.20	.20	.17	.21
Wald test against model		(1)	(2)	(2)	(2)	(7)	(2)
<i>F</i> statistics		37.22	14.56	5.63	2.31	21.16	4.60
Degrees of freedom		(15,1443)	(4,1443)	(4,1443)	(9,1443)	(11,1443)	(17,1443)
<i>p</i> value		.00	.00	.00	.01	.00	.00
Girls (<i>n</i> = 7,203)							
Slum	0.37** (0.07)	0.02 (0.08)	-0.14 (0.09)	-0.19 [†] (0.11)	-0.24 [†] (0.14)	-0.27 [†] (0.14)	-0.24 [†] (0.14)
Urban	0.45** (0.07)	0.16** (0.06)	0.01 (0.07)	-0.05 (0.09)	-0.04 (0.09)	-0.09 (0.09)	-0.06 (0.09)
<i>R</i> ²	.11	.21	.21	.21	.21	.19	.22
Adjusted <i>R</i> ²	.11	.20	.21	.20	.20	.19	.21
Wald test against model		(1)	(2)	(2)	(2)	(7)	(2)
<i>F</i> statistics		21.88	14.47	4.42	2.71	15.38	5.00
Degrees of freedom		(15,1430)	(4,1430)	(4,1430)	(9,1430)	(11,1430)	(17,1430)
<i>p</i> value		0.00	0.00	0.00	0.00	0.00	0.00

Notes: Figures are from a weighted ordinary least squares model, with robust standard errors clustered at PSU level shown in parentheses.

^a Age dummy variables for 4–7, 8–11, 12–17, 18–23, 24–35, 36–47, and 48–59 months old, with 0–3 as the excluded category.

^b Education dummy variables for mother and father. The dummy variables are for 1–4, 5–7, 8–9, 10–11, and 12 or more years of education.

^c Religion dummy variable for Muslim, with Hindu the excluded category. Caste dummy variables for each of scheduled caste, scheduled tribe, and other backward class, with *none of the above* the excluded category.

^d Wealth dummy variables for the household being in wealth category 2, 3, 4, and 5, respectively.

^e Area wealth is the percentage of households in wealth categories 2–5, calculated excluding the household itself.

^f Area health environment variables include water access, captured by the average time to fetch water and whether an improved source of drinking water is available, access to improved cooking fuel, sharing toilet with 10 or more households, access to improved toilet, and the average number of people per room. Each calculated as the average in PSU excluding the household itself.

[†] $p < .10$; * $p < .05$; ** $p < .01$

Table 3 Covariates of child health: Mortality

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
State and Survey Month Fixed Effects	No	Yes	Yes	Yes	Yes	Yes	Yes
Parental Education and Mother's Height ^a	No	Yes	Yes	Yes	Yes	No	Yes
Household Religion and Caste ^b	No	Yes	Yes	Yes	Yes	Yes	Yes
Household Wealth Status ^c	No	No	Yes	No	No	Yes	Yes
Area Wealth Distribution ^d	No	No	No	Yes	No	Yes	Yes
Area Health Environment ^e	No	No	No	No	Yes	Yes	Yes
All Children (<i>n</i> = 16,179)							
Slum	0.57** (0.09)	0.83 (0.15)	0.93 (0.17)	0.83 (0.17)	0.67 (0.18)	0.70 (0.19)	0.67 (0.18)
Urban	0.60** (0.06)	0.77* (0.08)	0.84 (0.09)	0.80 [†] (0.11)	0.72* (0.10)	0.75* (0.10)	0.73* (0.10)
Girl	1.12 (0.08)	1.11 (0.08)	1.11 (0.08)	1.11 (0.08)	1.11 (0.08)	1.11 (0.08)	1.11 (0.08)
Boys (<i>n</i> = 8,422)							
Slum	0.70* (0.13)	0.94 (0.19)	1.08 (0.23)	0.95 (0.23)	1.24 (0.34)	1.30 (0.35)	1.25 (0.34)
Urban	0.68** (0.08)	0.85 (0.11)	0.95 (0.13)	0.88 (0.16)	0.89 (0.16)	0.92 (0.16)	0.90 (0.16)
Girls (<i>n</i> = 7,757)							
Slum	0.45** (0.11)	0.68 (0.18)	0.75 (0.20)	0.69 (0.21)	0.30** (0.13)	0.32** (0.13)	0.31** (0.13)
Urban	0.52** (0.07)	0.70* (0.10)	0.74 [†] (0.11)	0.71 [†] (0.14)	0.54** (0.12)	0.56** (0.12)	0.55** (0.12)

Notes: Coefficients presented are hazard ratios from weighted Cox regressions, with robust standard errors clustered at PSU level shown in parentheses; a coefficient less than 1 indicates a lower risk of death compared with the reference group, whereas a coefficient greater than 1 indicates a higher risk compared with the reference group.

^a Education dummy variables for mother and father. The dummy variables are for 1–4, 5–7, 8–9, 10–11, and 12 or more years of education.

^b Religion dummy variable for Muslim, with Hindu the excluded category. Caste dummy variables for each of scheduled caste, scheduled tribe, and other backward class, with *none of the above* the excluded category.

^c Wealth dummy variables for the household being in wealth category 2, 3, 4, and 5, respectively.

^d Area wealth is the percentage of households in wealth categories 2–5, calculated excluding the household itself.

^e Area health environment variables include water access, captured by the average time to fetch water and whether an improved source of drinking water is available, access to improved cooking fuel, sharing toilet with 10 or more households, access to improved toilet, and the average number of people per room. Each calculated as the average in PSU excluding the household itself.

Table 4 Predicted stunting by area

	Percentage Stunted ^a			India 2011 ^b (millions)
	Pooled	Girls	Boys	
Rural	57.0	56.9	56.7	69.2
Urban	33.6	32.2	35.0	11.8
Slum	33.4	34.4	33.2	2.7
Urban	26.7	23.6	29.3	2.2
Rural	21.1	21.3	23.2	1.7

Notes: Predictions are based on Model 7, presented in Table 2 for height-for-age z score using the sample described earlier. These predictions should be taken as only suggested values.

^a All results use weighting provided in NFHS-3.

^b Calculated using the predicted percentage stunted and 2011 census information on number of children by area. According to the 2011 census, there were 121,285,762 rural children aged 0–6; 35,109,645 nonslum urban children; and 8,082,743 slum children, for a total child population of 164,478,150. This corresponds to 4.9 % of children living in slums, 21.4 % in urban nonslums, and 73.7 % in rural areas.