Sanitation externalities, disease, and children's anemia

Diane Coffey*

March 24, 2014

Working paper draft to be presented at PAA 2014.

Abstract

Anemia is a health problem with enormous economic consequences: it impairs cognitive ability, and reduces educational attainment and adult productivity. Globally, almost half of children have hemoglobin levels below the threshold for anemia. This paper uses three complementary empirical strategies to provide the first population-based evidence for the hypothesis that lack of sanitation, a public good with important externalities, contributes to a disease environment which causes hemoglobin deficiency. First, it finds a robust cross-country gradient between children's hemoglobin and lack of sanitation. Second, it shows that in India and Nepal, which both have poor sanitation coverage, children exposed to worse community sanitation have lower hemoglobin levels. Third, it shows that improvement in regional sanitation in Nepal between 2006 and 2011 predicts improvement in hemoglobin. Falsification tests and mechanism checks further suggest that the relationship is causal. In places where open defecation is widely practiced, policies to address anemia should put greater emphasis on improving sanitation, a public good with important disease externalities.

* *Affiliation*: Office of Population Research, Wallace Hall, Princeton University, Princeton, NJ 08540, coffey@princeton.edu. I appreciate comments and suggestions from Harold Alderman, Anjali Chikersal, Oliver Cummings, Angus Deaton, Michael Geruso, Jeffrey Hammer, H.P.S Sachdev, Dean Spears, and Belén Torondel. I thank participants of the Sanitation and Stunting Conference, held August, 2013 at the Delhi School of Economics, and at the National Food Security Act conference, held March, 2014 at Allahabad University, for helpful feedback. All errors are my own.

1 Introduction

Anemia is a disorder of the blood marked by lack of hemoglobin, the protein responsible for carrying oxygen throughout the body. Hemoglobin deficiency has serious health and economic consequences. In adults, it reduces work capacity and thus productivity, and is a major cause of maternal mortality in developing countries (Basta et al., 1979; Thomas et al., 2004; Rush, 2000). Children who are hemoglobin deficient have impaired cognitive ability and physical development, and are more susceptible to disease (Grantham-McGregor and Ani, 2001; Walter et al., 1989; Scrimshaw, 2000). Considering the high global prevalence of anemia (de Benoist et al., 2005), and the relationships between child cognitive ability, schooling, and adult productivity, anemia is a health problem with enormous economic consequences (Bloom and Canning, 2009; Alderman and Horton, 2007; Thomas and Frankenberg, 2002).

International variation in anemia rates is not well explained by differences in income; indeed, Alderman and Linnemayr (2009) suggest that economic growth alone is unlikely to reduce the high rates of anemia in developing countries. It is well known that, in sub-Saharan Africa, malaria is a major cause of anemia (Sachs and Malaney, 2002). In South Asia, however, where malaria is far less prevalent, explanations for high anemia rates have focused on dietary deficiencies (Pasricha et al., 2010; Yip and Ramakrishnan, 2002). Without denying the importance of food intake, this paper focuses on the disease environment, and particularly intestinal disease, as a cause of anemia. Although prior research has linked intestinal disease to nutritional outcomes (Smith et al., 2013; Checkley et al., 2008), this is the first paper to link poor sanitation environments with average hemoglobin levels in populations.

As in Spears (2013), this paper measures sanitation by constructing rates of open defecation. Open defecation is the practice of defecating outside without a toilet or latrine. Open defecation is practiced by over a billion people worldwide, and is especially common in South Asia; nearly 60 percent of the people in the world who defecate in the open live in India. Open defecation is a behavior of economic interest because it has significant negative externalities; even people who use toilets or latrines can be made sick by open defecation in their environments. Further, those who use toilets or latrines do not capture the full social benefit of the behavior.

This paper presents several empirical analyses of children's hemoglobin levels that indicate that open defecation contributes to a disease environment that causes anemia. It uses three complementary methodological approaches, including cross country variation, cross-sectional variation within South Asian countries, and over time variation within regions of Nepal to establish the robustness of the relationship between children's hemoglobin and the practice of open defecation. All of these analyses show a statistically significant and quantitatively important association between open defecation and children's hemoglobin levels. This paper also provides direct evidence for one of the proposed mechanisms leading from open defecation to anemia: it finds that in India, the association between taking deworming medicine and children's hemoglobin is lower for children who live in places with higher rates of open defecation, and thus have higher exposure to reinfection.

The approach of this paper is novel in that it applies econometric techniques to nationally representative and cross country data in order to answer questions that have mainly been studied by nutritionists. Studies by nutritionists about the causes of anemia tend to use smaller data sets than the ones used here, and, when presenting aggregate correlations, to use few control variables. Randomized trials of interventions to improve hemoglobin tend to focus on independent variables measured at the individual level, such whether a child received an iron supplement, fortified food, or a deworming pill. Because of the externalities associated with open defecation, it is most appropriately studied at the community level. Since the data for medical and epidemiological studies are often collected within communities, they cannot be used to study the effects of community level variables.

The paper proceeds as follows: section 2 presents background information about the causes of anemia, the measurement of sanitation externalities, and links between sanitation and nutritional outcomes. Section 3 describes the data used in the empirical analyses. The paper then provides three complementary analyses; section 4 estimates a cross-country hemoglobin-open defecation gradient; section 5 estimates the relationship between a child's hemoglobin and her community's level of open defecation in India and in Nepal; section 6 shows a robust relationship between regional reduction in open defecation in Nepal between 2006 and 2011 and improvement in hemoglobin status of children. Section 7 presents direct evidence for one of the main biological mechanisms leading from open defecation to anemia; it finds a weaker association between taking deworming medicine and hemoglobin levels among children living in communities with high levels of open defecation. Section 8 discusses these analyses and their implications for policy. These results, combined with traditional economic justifications for investment in public goods, suggest that policy makers and program implementers should consider sanitation provision an important part of the public response to anemia, especially where open defecation is widely practiced.

2 Background

2.1 Causes of hemoglobin deficiency

Hemoglobin is a protein which resides in red blood cells, and which binds to iron in order to attract oxygen and carry it throughout the body. There are several causes of low hemoglobin, these involve either too little production or too much destruction of hemoglobin. A major cause of low hemoglobin production is iron deficiency. However, low hemoglobin can can also be caused by lack of vitamin B12 and folic acid, two nutrients necessary for the production of red blood cells.

Intestinal parasites are well known to cause of anemia by causing blood loss in the stool, lack of appetite, increased motility of food through the intestine, competition for nutrients, and damage to the intestinal wall that leads to decreased absorption of nutrients, including iron, vitamin B12 and folic acid (Rosenberg and Bowman, 1982). Malaria is an important cause of anemia in sub-Saharan Africa. Malaria is a parasite which attacks red blood cells, which are in turn attacked by the host's immune system.

Environmental enteropathy, known as tropical sprue in an older medical literature, is a disease which alters the lining of the intestine and inhibits absorption of calories and nutrients. It is believed to be caused by the ingestion of large quantities of fecal pathogens (Walker, 2003; Humphrey, 2009; Lin et al., 2013; Kosek et al., 2013), and it has been documented that enteropathy affects the absorption of vitamin B12 and folic acid, two essential nutrients for the production of hemoglobin (Nath, 2005).

Poor diets, particularly among young children, are another often-cited cause of anemia in developing countries (Yip and Ramakrishnan, 2002; Tolentino and Friedman, 2007). Particularly in South Asia, the introduction of solid foods is delayed in comparison to WHO recommendations, and children are fed small quantities of food that contain inadequate amounts of essential nutrients (Menon, 2012).[1] However, it is often difficult to correct hemoglobin deficiency with dietary supplementation; Friis et al. (2003) find that, in Kenyan school children, a single dose of deworming medicine was as effective in improving hemoglobin levels as a daily supplement of 13 micronutrients (including iron, vitamin B12 and folate) taken for eight months. This suggests that an intervention aimed at combatting infectious disease can be as effective at improving hemoglobin as a more laborious dietary intervention.

2.2 Measuring sanitation externalities

Open defecation, without a toilet or latrine, is a practice that occurs in much of the developing world, but it is especially common in South Asia. Figure 1 shows the UNICEF & WHO (2012) estimates of regional rates of individual open defecation, as well as the author's calculations of household level open defecation rates for India and Nepal from the Demographic and Health Surveys. It shows that open defecation was practiced by

[1] A strict vegetarian diet, containing no animal proteins, would be lacking in vitamin B12.

about 25% of the population of sub-Saharan Africa and over 40% of the population of South Asia in 2010, and by about 55% of households in India in 2005.

Open defecation rates are a useful measure of the sanitary environment to which a child is exposed. The greater the fraction of households which do not use a toilet or latrine, the greater the chance that a child comes into contact with fecal germs or fecally transmitted parasites. Figure 2 provides a simple illustration of the negative externalities associated with open defecation. It plots the hemoglobin-community open defecation gradient separately for Indian children whose households do and do not have latrines: there is a clear negative slope for both groups of children.

For the cross country analysis, it is possible to create a measure of the density of open defecation by matching data on population density with open defecation rates. This measure takes into account that open defecation will have more of an impact on the disease environment where it is practiced in high population density areas. In high density areas, children are more likely to come into contact with fecally transmitted infections than they would be in low density areas.

2.3 Links between sanitation and nutritional outcomes

This is not the first paper to document a relationships between sanitation and nutritional outcomes. Recent work in economics and epidemiology suggests that open defecation plays an important role in child stunting, that is, in inhibiting the physical growth of children. Spears (2013) shows that open defecation per square kilometer explains 65 percent of international variation in child height, and that India's higher prevalence of open defecation explains the entire India-Africa child height gap. Spears and Hammer (2012) report on a randomized experiment in Maharastra, India that found an effect of sanitation improvement on children's height, and Cameron et al. (2013) present a randomized experiment in Indonesia that shows that a sanitation intervention had a positive effect on height. Finally, Lin et al. (2013) find that Bangladeshi children growing up in environments less contaminated by fecal pathogens grow taller than children growing up in more contaminated environments.

This paper proposes that open defecation causes hemoglobin deficiency through the transmission of intestinal parasites and possibly also environmental enteropathy.[2] It has long been known that open defecation spreads intestinal parasites; Cairncross (2003) cites research from the 1930s that describes how variation in community latrine use in the southern United States predicted parasite infections in children living in different communities.[3] Although the link between open defecation and enteropathy less well understood, it is hypothesized that open defecation exposes people to the kinds of bacteria that, when ingested in large quantities, leads to decreased absorption of micronutrients necessary for the production of hemoglobin (see Walker (2003), Nath (2005) and Humphrey (2009)). Medical researchers have hypothesized a link between enteropathy and anemia as long ago as the 1920s (Baumgartner and Smith, 1927).

3 Data

The data used for these analyses are primarily from the Demographic and Health Surveys (DHS). These nationally representative surveys of poor and middle income countries are publicly available.[4] For the cross country analysis in section 4, I use all of the 81 DHS surveys that, according to the DHS' Statcompiler (www.statcompiler.com), have ever collected data on children's hemoglobin. These surveys are from different 45 countries and 18 different years, spanning 1995-2012. See the appendix for a list of countries and years included in this analysis. For the cross sectional analysis in section 5, I use Nepal's 2011 DHS and India's 2005 DHS. For the fixed effects analysis in section 6, I use Nepal's 2006 and 2011 surveys.[5]

[2] It is also possible that maternal disease caused by open defecation could lead her to pass low iron stores on to her baby. See Allen (1997).

[3] Children infected with parasites pass them through their bowel movements to the soil; many parasite eggs can live in the soil until they come into contact with new hosts, such as other children who come to the same places to defecate in the open.

[4] For more information, visit www.measuredhs.com.

[5] Unfortunately, India's 2005 DHS did not report district identifiers, so it is not possible to do a similar analysis with the two Indian surveys that collected hemoglobin.

In this section, I will discuss the measurement of the two main variables, hemoglobin and open defecation. Section 4.1 details the control variables used in the cross country analysis; section 5.2 discusses the controls for the cross sectional analysis; and section 6.1 discusses controls for the fixed effects analysis.

Hemoglobin. The DHS measures hemoglobin using the HemoCue®method, in which a surveyor introduces a drop of blood from the respondent's finger into a portable device which reports the respondent's hemoglobin level in the field.[6] Instead of collapsing the hemoglobin data into anemia categories, I instead use the entire range of hemoglobin data. Therefore, the results are not sensitive to the particular cutoff for anemia used, and the power of the analysis is greater than it would be if categories for the severity of anemia were used.

Open defecation. Variables for open defecation are constructed at the country level, the regional level, or at the community (primary sampling unit, or PSU) level, depending on the analysis. In all cases, household level information on regular defecation behavior was used to construct rates of open defecation. For instance, community level open defecation is the fraction of households in the PSU that report open defecation without a toilet or latrine. The same procedure was applied to region level and country level open defecation. To account for differential probability of selection into the DHS sample, the household sampling weight was used when calculating the fraction of households which defecate in the open in a place.

For the analysis that uses density of open defecation as an independent variable in section 4, each of the country years was matched with time invariant data from the World Bank on land area (see World Bank (2013)) and with time variant data from the Penn World Tables on population (see Heston et al. (2012)). These data were used to construct an estimate of the number of people who defecate in the open per square kilometer. The estimate of open defecation per square kilometer is derived by multiplying the fraction

[6]See Measure DHS (2013) for more information about collection of hemoglobin data in the field. Kapoor et al. (2002) compares the HemoCue®method to other methods of testing hemoglobin in India. This paper uses the hemoglobin measurement that is unadjusted for altitude.

of households that defecate in the open by the population of the country in the year of interest, dividing by the land area in kilometers, and taking the log.[7]

4 Evidence from international comparisons

Panel A of figure 3 shows the relationship between the average hemoglobin level of children between the ages of 6 and 35 months in a particular country and year and the open defecation rate in that country in that year. The size of the circle for each observation is proportional to the size of the population of that country in that year. It shows a clear negative relationship between sanitation and hemoglobin levels of children. Panel B of figure 3 shows a similar graph, except that it uses the density of open defecation as the independent variable. Panel C of figure 3 shows the same relationship net of malaria incidence given in Korenromp (2005). Panel C suggests that controlling for malaria incidence makes the relationship between hemoglobin and open defecation stronger; the observations are more clustered around the regression line. This is similarly true of panel D, which shows the relationship between hemoglobin, net of malaria incidence, and density of open defecation, net of malaria incidence. This is because many of the sub-Saharan African countries with high rates of malaria have relatively low rates of open defecation. Therefore, net of malaria incidence, their hemoglobin levels are closer to what their open defecation rates would predict.

4.1 Empirical strategy

Table 1 presents the results of regressions of the average hemoglobin level of children in a given country year on the density of open defecation in that country and year (panel A) and on the fraction of households defecating in the open in that country and year (panel B). It also includes several control variables. The regression equation that corresponds

[7]To the extent that households which defecate in the open are larger than households which practice sanitary feces disposal, this will be an underestimate of exposure to fecally transmitted diseases.

with the regression in column 4 of panel A of table 1 is

$$h_{cry} \quad = \quad \beta ln(od_{cry}) + \gamma m_{cr2004} + \theta GDP_{cry} + \delta_r + \varepsilon_{cry}. \tag{1}$$

h_{cry} is the average hemoglobin level of children in country c, in region r, in year y. $ln(od_{cry})$ is the log of the estimated number of people who defecate in the open per square kilometer in that country and year. The controls are:

- m_{cr2004}: Malaria: The control for malaria exposure is a time invariant estimate of malaria incidence in children under five generated by Korenromp (2005) for the WHO. For Africa, these estimates were generated based on 22 longitudinal studies from populations with no access to malaria prevention. The studies defined malaria as fever with malaria parasitemia on a blood slide. In some countries in southern Africa, national malaria case notification rates were used. Korenromp (2005) points out that for countries outside of Africa, there were fewer longitudinal studies on which to base the malaria estimates.

- GDP_{cry}: GDP per capita: The control for GDP per capita is time variant and is taken from the Penn World Tables. GDP per capita is in US dollars, converted using the Laspeyere's PPP conversion (Heston et al., 2012).

- δ_r: Region fixed effect.

4.2 Results

Table 1 adds controls in stages. The coefficient in the first column, which presents the simple association between average hemoglobin and log of open defecation per square kilometer, suggests that a doubling of open defecation per square kilometer is associated with a 0.14 gm/dL decrease in the average children's hemoglobin level. The magnitude of this coefficient is fairly stable; if anything the magnitude increases slightly as controls are added. Column 2 adds the malaria control, and column 3 adds a control for GDP per capita. Column 4 adds a region fixed effect, and column 5 adds a year fixed effect. Due to the small number of observations in each region and the small number of observations for many of the years, the results are not robust to the inclusion of both region and year fixed effects.

Panel B repeats these regressions using the fraction of households in the country which defecate in the open as the independent variable. In the controlled regressions presented in columns 3 and 8, the coefficients on open defecation suggest that compared with a country year in which everyone has access to sanitation, in a country year in which there is no sanitation coverage, the average of children's hemoglobin would be 0.7 gm/dL lower.

As in Alderman and Linnemayr (2009), this study also finds that the relationship between hemoglobin and GDP is weak, and further, that it is not statistically significant when other variables are added to the regression. In all specifications, malaria incidence is a strong predictor of average hemoglobin. The magnitude of the coefficient on malaria incidence at first appears larger than the magnitude of the coefficient on open defecation. However, the range of open defecation is much greater than the range of malaria incidence. The coefficients in column 3 of panel B suggest that the difference in predicted hemoglobin associated with being a country with 75th rather than 25th percentile of open defecation—0.25 gm/dL—is about the same as the difference in predicted hemoglobin associated with being a 75th rather than 25th percentile country in terms of malaria incidence.

Table 1 also reports the R^2 associated with the different specifications. Open defecation per square kilometer by itself explains 26.3 percent of the variation in average hemoglobin levels across countries.

5 Evidence from India & Nepal

The results presented in section 4 suggest that across countries, and across time, open defecation predicts average hemoglobin levels. However, in a small sample of country years it is difficult to control for the many other factors that also influence hemoglobin, and which could potentially contribute to an overestimate of the open defecation-hemoglobin gradient. The individual level analysis allows controls for a number of potentially important predictors of hemoglobin, such as food intake, medicine, and measures of household and community socioeconomic status.

In this section, I present individual level results from from India and Nepal. These two countries were chosen because they have high levels of open defecation and very little malaria. Open defecation is widespread, particularly in India, and its practice is less of a marker of community deprivation than in other settings. Indeed, Geruso and Spears (2014) find that Muslim communities in India are less likely to practice open defecation in India than Hindu communities, despite their relatively lower socioeconomic status.[8]

5.1 Summary statistics

Table 2 presents child level summary statistics about hemoglobin and early life health in India and Nepal. Nepali children have higher hemoglobin levels for every age group than Indian children. Figure 4 shows the age pattern of hemoglobin concentration for India and Nepal. The figure suggests that Indian children experience a steep decline in hemoglobin levels between 6 and 18 months of age that is less pronounced in Nepali children. This is consistent with the poor complementary feeding practices that are prevalent in India; by six months of age breastmilk is no longer nutritionally adequate for growing children, but many families in India do not supplement their children's diets with complementary food until they are older. Table 2 shows that in India, children are fed solid or semi-solid food fewer times in the last 24 hours than in Nepal. Table 2 also shows that Indian children are also less likely to have had parasite medicine in the six months prior to the survey, and are less likely to be taking iron supplements.

Table 3 presents summary statistics about PSU level open defecation. In the average rural PSU in the India, 2005 survey, 61 percent of households defecated in the open. This

[8]Geruso and Spears (2014) show that open defecation rates can explain the paradoxical survival advantage of Muslim infants described by Bhalotra et al. (2010).

figure was 11 percent for urban PSUs. In the mean rural PSU in the Nepal, 2011 survey, 40 percent of households defecated in the open, compared with 12 percent of urban PSUs.

5.2 Empirical strategy

Panel A of table 4 shows the results of regressions of children's hemoglobin on community level open defecation and food and medicine controls. It shows the results regressions separately for India and Nepal. Panel B uses a similar strategy but includes socioeconomic and demographic controls as well.

The full specification, for the India sample, is shown in column 3 of panel B of table 4. The regression equation that corresponds with column 3 of panel B for India is

$$\begin{aligned} h_{ip} = \; & \beta od_p + M_{ip}\Theta + F_{ip}\Lambda + \gamma elec_p + \delta urban_p + E_{ip}\Phi + SES_{ip}\Psi + \\ & household\,toilet_{ip} + (A_{ip} \times sex_{ip})\Pi + \varepsilon_{ip}. \end{aligned} \tag{2}$$

h_{ip} is the hemoglobin level of child i in PSU p. β, the coefficient on od_p, the fraction of households in the child's PSU which practice open defecation, is the coefficient of interest. The controls are introduced in groups.

- M_{ip}: Medicine: indicators for having taken parasite medicine in the six months prior to the survey and for taking iron supplements.

- F_{ip}: Food: indicator for breast feeding, the number of times the child was fed solid/semi-solid food in last 24 hours, whether the child ate meat in last 24 hours, and whether she/he ate dark green leafy vegetables in last 24 hours. In the Demographic and Health Surveys, these variables were only recorded for the youngest child in each household. Hence, there is a drop in sample size from column 2 to column 3, and children from smaller households are over represented in column 3.

- $elec_p$: Electrification: the fraction of households in a child's PSU with electricity. This is intended to control for the overall level of development of the child's PSU.

- $urban_p$: Urban: an indicator for whether the child's PSU is urban.

- E_{ip}: Mother's education: in column 1 of panel B, a linear term for mother's years of education, in columns 2 and 3 of panel B, indicators for mother's years of education.

- SES_{ip}: Household economic status: in column 1 of panel B, a linear term for the percentile of the DHS wealth index to which the child's household belongs, columns 2 and 3 of panel B include dummy variables for whether or not the child's household owns a radio, a TV, a fridge, a bicycle, a motorcycle, a car, and whether or not it has electricity. It also includes indicators for the type of drinking water the child's house uses (17 options), what type of floor material the child's house has (12 options), as well as what kind of wall (15 options) and roof material (14 options) the house has.

- $household\, toilet_{ip}$: Household toilet: for India, this is a dummy variable for whether the household has its own toilet. Because a household's own defecation behavior is highly correlated with that of its neighbors, we would expect the inclusion of this control to lead to a reduction in the magnitude of the hemoglobin-open defecation gradient. In India, where there are many households per PSU, this control is included to illustrate that community open defecation matters for hemoglobin above and beyond a household's own defecation behavior.

- $A_{ip} \times sex_{ip}$: Sex-specific age in months indicators: these are 108 indicators for the age in months of the child (from 6 to 59 months), interacted with the sex of the child. These indicators control for the age and sex patterns of hemoglobin status.

5.3 Results

Panel A of table 4 shows the results of regressions of individual hemoglobin on community open defecation with controls for food intake and medicine. The literature suggests that the associations between hemoglobin and iron-rich food intake, as well as between hemoglobin and iron supplementation or deworming medicine, should be positive and statistically significant. The India sample of panel A of table 4 shows the expected relationships. The coefficients with the largest magnitudes correspond with whether the child has taken parasite medicine in the six months before the survey and whether the child was fed meat in the 24 hours before the survey. The coefficients on food and medicine are not statistically significant in Nepal, but this may be because the sample size is much smaller; the Nepal sample is less than a twenty-fifth the size of the India sample. This may also be because in general, child care and feeding practices are better in Nepal than in India (as shown in table 2), so that small differences in care and feeding are not associated with the same improvements in hemoglobin in Nepal as in India.

Across the specifications in panel A, the coefficient on open defecation is remarkably similar. Compared to growing up in a community with no open defecation, growing up in a community with 100 percent open defecation is associated with having about 0.6 gm/dL less hemoglobin, which is enough to move the mean 24-59 month old child in Nepal from being anemic to not anemic. In India, where hemoglobin levels are lower, this would not be the case, though it would represent nearly a 0.4 standard deviation difference in hemoglobin levels among children 6-59 months.

Panel B adds a variety of socioeconomic and demographic controls. Column 1 of panel B includes a linear term for mother's education, for the child's household's wealth percentile, for whether the child lives in an urban household, and for the fraction of households in her PSU which have electricity, which is meant to proxy for the overall level of development or infrastructure in the child's community. In both India and Nepal, the coefficient on community level open defecation is robust to these controls.

Column 2 of panel B introduces an additional set of controls about the child's household. In addition to controlling for PSU level electrification and urban residence, it includes an indicator for each year of mother's education, and 65 dummy variables for asset ownership, housing type and water source. For India, it also includes a control for whether or not the household has its own toilet. Column 3 of panel B adds in the medicine and food controls that were shown in panel A.

In columns 2 and 3 of panel B of the India sample, because of the control for *own toilet*, the coefficient on open defecation represents the association between hemoglobin and community open defecation net of the household's own defecation practices, and of the other controls included in the regression. This specification probably underestimates the hemoglobin-open defecation gradient because of the high correlation between household and community defecation behavior. Nonetheless, this specification is useful in that it shows that, neighbors' defecation behavior is associated hemoglobin at least as strongly as household behavior. Indeed, the magnitude of the coefficient on community level open defecation is more than twice the magnitude of the coefficient on *own toilet*. This control is not included in the Nepal regressions since there is a much smaller number of households per PSU, so the sanitary practices of the child's household and the practices of other households in the PSU are highly collinear.

The coefficients on community level open defecation are statistically significant in all of the specifications. The robust gradients described in these regressions suggest a causal effect of community sanitation on hemoglobin, but it is important not to interpret the magnitudes of the coefficients literally. The coefficient in column 1 of panel A, from a regression that contains no controls, is likely an overestimate of the "true effect" of sanitation on hemoglobin. This is because, to the extent that open defecation is correlated with poverty or bad feeding practices, the association between open defecation and hemoglobin will pick up these relationships as well. But the coefficients on the heavily controlled regressions are likely an underestimate, since what looks like an effect of perhaps, wealth, may have less to do with wealth than the fact that poorer people are, on average, more likely to defecate in the open, and to have neighbors who defecate in the open.

6 Evidence from change over time within Nepali regions

This analysis looks at the relationship between hemoglobin and open defecation in Nepal using region fixed effects. This strategy controls for everything that is time invariant about a region. Nepal was chosen for this analysis because it had two DHS in the recent past, both of which measured hemoglobin and provided region identifiers.[9] It also has high levels of open defecation and a significant change in open defecation; nationally, there was a 15 percentage point reduction in open defecation between 2006 and 2011, from 50 percent to 35 percent. Table 3 reports the regional improvements in open defecation over this period.

6.1 Empirical strategy

Table 5 presents the results of regressions using two different fixed effects strategies. In columns 1-3, individual hemoglobin levels are regressed on open defecation, collapsed to the regional level, and region fixed effects, as well as household and individual level controls. The regression equation that corresponds to column 3 of table 5 is

$$\begin{aligned} h_{iyr} = \quad & \beta od_{yr} + \gamma elec_{yr} + \alpha_y + \delta_r + E_{iyr}\Phi \\ & + SES_{iyr}\Psi + (A_{iyr} \times sex_{iyr})\Pi + \varepsilon_{iyr}. \end{aligned} \tag{3}$$

h_{iyr} is the hemoglobin level of child i, in year y, in region r. A region is defined as the rural or urban part of each of 13 regions identified by the DHS. One region, the western mountain, has only rural households, and no urban households, so the regression includes fixed effects for 25 regions. od_{yr} is the fraction of households that defecate in the open in the region in that year. $elec_{yr}$ is the fraction of households which have electricity in that region in that year. α_y is an indicator for the year 2011, and δ_r is a region fixed effect.

[9] Unfortunately, the Indian survey does not include district identifiers in 2005, so it is not possible compare change in district level open defecation with improvement in district level hemoglobin levels.

E_{iyr} is a vector of indicators for mother's years of education, SES_{iyr} includes the same 65 dummy variables for assets, water source and housing materials as described above. The sex specific age-in-months indicators are also included.

In columns 1-3, the coefficient on regional open defecation identifies the association between change in open defecation in a region and the average improvement in children's hemoglobin in that region. Over time, rather than cross-place variation in open defecation is used to identify the association. An advantage of this strategy is that it controls for everything about the region that does not change over time. A disadvantage of this strategy is that region level open defecation is probably an imprecise measure of a child's exposure to fecally-transmitted diseases; we would not expect the open defecation of a household living in one part of the region to affect the hemoglobin status of a child living in another part, many miles away. Community level open defecation would be a more accurate measure of a child's exposure to fecally-transmitted diseases.

Therefore, for columns 4-6, the empirical strategy uses both across time and across place variation to identify the association between open defecation and hemoglobin. Instead of using od_{yr} as the measure of open defecation, it uses od_{yrp}, which is the fraction of households that defecate in the open in year y, region r, and PSU p.

6.2 Results

The results reported in columns 1-3 of table 5 show that improvement in open defecation in a region predicts improvement in the hemoglobin levels of children. These specifications control for everything that is constant about a region across time, as well as for characteristics of the children and the households. The coefficient is similar in all three specifications, and suggests that a reduction of regional open defecation from 100 percent to 0 percent is associated with an improvement in average hemoglobin of 1.4 gm/dL. This is a large improvement which corresponds to a one standard deviation increase in 2011 hemoglobin levels. Of course, the mean reduction in open defecation in the regions of Nepal was much less than 100 percent; the median region reduced open defecation by 14 percentage points. The associated change in hemoglobin was about 0.2 gm/dL.

The results reported in columns 4-6 use cross sectional as well as cross time variation to identify the association between open defecation and hemoglobin. The coefficients on od_{yrp} are statistically significant in all of the specifications.

7 Mechanism check: Interaction of open defecation with parasite medicine

Is there any direct evidence for the role of open defecation in creating a disease environment that leads to poor hemoglobin status? Table 6 uses data from the India, 2005 survey to show the results of regressions of hemoglobin on community level open defecation, parasite medicine, and the interaction of parasite medicine and open defecation. The coefficient on the interaction is negative and statistically significant. This suggests that in places with more open defecation, taking parasite medicine may be less effective in improving hemoglobin levels.

This finding is consistent with the theory presented in section 2.3 that open defecation spreads parasites that cause hemoglobin deficiency. In places with more open defecation, we would expect parasite medicine to be less effective because the chance of reinfection is higher. Indeed, the coefficient on the interaction of open defecation and parasite medicine in column 1 of table 6 suggests that taking parasite medicine has less than half the association with hemoglobin where community open defecation is 100 percent rather than 0 percent.

This interaction is not simply due to poor feeding habits in communities with lots of open defecation; column 2 of table 6 shows that the interaction of open defecation and parasite medicine is statistically significant even when controls for food intake are added. Column 3 shows that the coefficient on the interaction is also robust to the inclusion of socioeconomic status controls.

8 Discussion & policy implications

This section discusses three questions. First, should the coefficients on open defecation in these analyses be interpreted as causal? Second, is the magnitude of the gradient between open defecation and hemoglobin plausible, and how does it compare with the effects of randomized interventions designed to fight anemia? Third, what are the implications of this study for policy?

These analyses suggest that the relationship between open defecation and children's hemoglobin is not due to associations with food intake, wealth or maternal education. Nor is it driven by an association between community open defecation and the overall level of development in the community. Additionally, this paper has presented direct evidence for the role of open defecation in spreading intestinal parasite infection, which causes anemia.

An observational study of the relationship between sanitation and hemoglobin may be less likely to overestimate the relationship than would be similar studies of hemoglobin and other independent variables. One thing that distinguishes the interpretation of the magnitude of the coefficient on community open defecation from the interpretations of coefficients on for instance, the food and medicine variables, is that whether or not a household uses a latrine is probably not a response to parents' perceptions of the health or hemoglobin status of the child. It is plausible that parents feed their child in response to the child's perceived health status. If that is the case, children showing signs of anemia might be more likely to be fed supplements, medicine or high iron foods than kids who were not anemic. This, combined with evidence that dietary interventions work better for anemic children than for the average child (Taylor-Robinson et al., 2012), suggests that the coefficients on these variables could be overestimates of their effects on the average child. Sanitation, by contrast, is a public good, which is harder to obtain than medicine or healthy food. It involves fixed costs of construction, as well as maintenance costs. Additionally, parents in poor countries are largely unaware of the negative health impacts of open defecation (Jenkins and Curtis, 2005).

To my knowledge, there have unfortunately been no randomized controlled trials of the effect of latrine or toilet provision on hemoglobin status.[10] This study suggests that such a trial would make an important contribution to the literature on the determinants of anemia in poor countries, and particularly in South Asia. It is important to note that two randomized trials, one from India and one from Indonesia, have found an effect of sanitation programs on child height (Spears and Hammer, 2012; Cameron et al., 2013), and that the biological mechanisms linking open defecation and height are similar to those proposed to link open defecation and hemoglobin status.

Are this study's estimates of the magnitude of the relationship between open defecation and hemoglobin plausible? From the Indian data, the difference in hemoglobin associated with living in an open defecation free locality vs. a locality in which every defecates in the open is about 0.6 gm/dL of hemoglobin. Interventions such as micronutrient supplementation (Friis et al., 2003), iron supplementation (Lind et al., 2003), fortification (Van Stuijvenberg et al., 1999), and treating for parasites (Friis et al., 2003; Taylor-Robinson et al., 2012) have effect sizes that range from 0.2 gm/dL to 0.4 gm/dL. In comparison, the sanitation-hemoglobin gradient is quite large, though not implausibly so. It is important to note that the gradients estimated in this study describe the difference in hemoglobin status associated with no vs. full sanitation coverage. In communities that already have some sanitation coverage, the improvement in hemoglobin associated with improving sanitation would probably be less. Additionally, here, as in randomized trials, the association between hemoglobin and the independent variable depends on context. For instance, open defecation is probably more detrimental to hemoglobin where population density is high, deworming medicines may work better where reinfection is less likely, and iron supplementation may be more useful where diets include adequate amounts of other micronutrients needed for hemoglobin production.

The results of this study suggest that open defecation likely has a quantitatively

[10] The SHINE trial, currently underway in Zimbabwe, aims to demonstrate the link between the sanitation environment, environmental enteropathy, and anemia. See Humphrey (2014).

important effect on children's hemoglobin status. Such research is important because researchers and policy makers tend to consider anemia and other nutritional diseases to be problems of inadequate food intake, and to overlook the important role of disease in determining nutritional outcomes. Indeed, recommendations and interventions by leading development organizations almost always focus on food intake based interventions, either in the forms of iron, vitamin B12 and folate supplementation or fortification, or through efforts to encourage people to diversify their diets.

The emerging understanding, to which this paper adds, of the importance of poor sanitation in creating a disease environment that adversely affects nutritional outcomes suggests that, particularly in South Asia, policy makers should consider sanitation an important priority. Although development organizations such as the WHO and UNICEF have included poor hygiene and sanitation among the contributing factors to anemia in young children (see WHO et al. (2001) and UNICEF (2002)), the elimination of open defecation is rarely among priority policy recommendations or the focus of programs implemented to fight anemia. The case for making sanitation a public policy priority rests not only on the strength of relationships between sanitation and nutritional outcomes, but also on the negative externalities associated with poor sanitaiton. This is particularly true in South Asia, and specifically in India, where open defecation is widespread and nutritional indicators are abysmal.

References

Alderman, H. and S. Horton (2007). The economics of addressing nutritional anemia. *Nutritional anemia* (19).

Alderman, H. and S. Linnemayr (2009). Anemia in low-income countries is unlikely to be addressed by economic development without additional programs. *Food & Nutrition Bulletin 30*(3), 265–269.

Allen, L. (1997). Pregnancy and iron deficiency: unresolved issues. *Nutrition reviews 55*(4), 91–101.

Basta, S., Soekirman, D. Karyadi, and N. Schrimshaw (1979). Iron deficiency anemia and the productivity of adult males in Indonesia. *American Journal of Clinical Nutrition 32*, 916–925.

Baumgartner, E. and G. Smith (1927). Pernicious anemia and tropical sprue. *Archives of Internal Medicine 40*(2), 203–215.

Bhalotra, S., C. Valente, and A. Van Soest (2010). The puzzle of muslim advantage in child survival in india. *Journal of health economics 29*(2), 191–204.

Bloom, D. E. and D. Canning (2009). Population health and economic growth. *Health and Growth*, 53.

Cairncross, S. (2003). Sanitation in the developing world: current status and future solutions. *International Journal of Environmental Health Research 13*(S1), S123–S131.

Cameron, L., S. Olivia, and M. Shah (2013). *Impact evaluation of a large-scale rural sanitation project in Indonesia.* World Bank.

Checkley, W., G. Buckley, R. H. Gilman, A. M. Assis, R. L. Guerrant, S. S. Morris, K. Mølbak, P. Valentiner-Branth, C. F. Lanata, R. E. Black, et al. (2008). Multi-country analysis of the effects of diarrhoea on childhood stunting. *International journal of epidemiology 37*(4), 816–830.

de Benoist, B., E. McLean, I. Egli, and M. Cogswell (2005). Worldwide prevalence of anaemia 19932005: WHO global database on anaemia. Technical report, World Health Organization and Centers for Disease Control and Prevention.

Friis, H., D. Mwaniki, B. Omondi, E. Muniu, F. Thiong'o, J. Ouma, P. Magnussen, P. Geissler, and K. F. Michaelsen (2003). Effects on haemoglobin of multi-micronutrient supplementation and multi-helminth chemotherapy: a randomized, controlled trial in kenyan school children. *European journal of clinical nutrition 57*(4), 573–579.

Geruso, M. and D. Spears (2014). Sanitation and health externalities: Resolving the muslim mortality paradox. working paper.

Grantham-McGregor, S. and C. Ani (2001). Review of studies on the effect of iron deficiency on cognitive development in children. *The Journal of Nutrition 131*, 649S668S.

Heston, A., R. Summers, and B. Aten (2012). Penn world table version 7.1.

Humphrey, J. (2009). Child undernutrition, tropical enteropathy, toilets, and handwashing. *Lancet 374*, 10321035.

Humphrey, J. (2014). Sanitation, hygiene, infant nutrition efficacy project.

Jenkins, M. and V. Curtis (2005). Acheiving the 'good life': Why some people in rural Benin want latrines. *Social Science and Medicine 61*, 24462459.

Kapoor, S., U. Kapil, S. N. Dwivedi, K. Anand, P. Pathak, and P. Singh (2002). Comparison of hemocue method with cyanmethemoglobin method for estimation of hemoglobin. *Indian pediatrics 39*(8), 743–746.

Korenromp, E. (2005). Malaria incidence estimates at country level for the year 2004. proposed estimates and draft report. *World Health Organization, Roll Back Malaria. Geneva: WHO*.

Kosek, M., R. Haque, A. Lima, S. Babji, S. Shrestha, S. Qureshi, S. Amidou, E. Mduma, G. Lee, P. P. Yori, et al. (2013). Fecal markers of intestinal inflammation and permeability associated with the subsequent acquisition of linear growth deficits in infants. *The American journal of tropical medicine and hygiene 88*(2), 390.

Lin, A., B. Arnold, S. Afreen, R. Goto, T. M. N. Huda, R. Haque, R. Raqib, L. Unicomb, T. Ahmed, J. Colford, and S. Luby (2013). Household environmental conditions are associated with enteropathy and impaired growth in rural Bangladesh. *The American Journal of Tropical Medicine and Hygiene 12*.

Lind, T., B. Lönnerdal, H. Stenlund, D. Ismail, R. Seswandhana, E.-C. Ekström, and L.-Å. Persson (2003). A community-based randomized controlled trial of iron and zinc supplementation in indonesian infants: Interactions between iron and zinc. *The American journal of clinical nutrition 77*(4), 883–890.

Measure DHS (2013). Malaria indicator survey: Anemia and malaria field testingmanual. Technical report, Measure DHS, Calverton, MD. Available: www. malariasurveys. org/toolkit.

Menon, P. (2012). The crisis of poor complementary feeding in South Asia: Where next? *Maternal & Child Nutrition 8*(s1), 1–4.

Nath, S. (2005). Tropical sprue. *Current Gastroenterology Reports 7*, 343–349.

Pasricha, S.-R., J. Black, S. Muthayya, A. Shet, V. Bhat, S. Nagaraj, N. Prashanth, H. Sudarshan, B.-A. Biggs, and A. S. Shet (2010). Determinants of anemia among young children in rural india. *Pediatrics 126*(1), e140–e149.

Rosenberg, I. and B. Bowman (1982). Intestinal physiology and parasitic diseases. *Reviews of Infectious Diseases 4*(4), 763–767.

Rush, D. (2000). Nutrition and maternal mortality in the developing world. *American Journal of Clinical Nutrition 72*(suppl), 212S–240S.

Sachs, J. and P. Malaney (2002). The economic and social burden of malaria. *Nature 415*(6872), 680–685.

Scrimshaw, N. (2000). *Infection and Nutrition: Synergistic Interactions*, Chapter VI.3. Cambridge University Press.

Smith, M. I., T. Yatsunenko, M. J. Manary, I. Trehan, R. Mkakosya, J. Cheng, A. L. Kau, S. S. Rich, P. Concannon, J. C. Mychaleckyj, et al. (2013). Gut microbiomes of malawian twin pairs discordant for kwashiorkor. *Science (New York, NY) 339*(6119), 548–554.

Spears, D. (2013). Sanitation and open defecation explain international variation in childrens height: Evidence from 140 nationally representative household surveys. Technical report, RICE working paper. Available: www. riceinstitute. org.

Spears, D. and J. Hammer (2012). Village sanitation externalities and children's human capital: Evidence from a randomized experiment by the Maharashtra government. Technical report, RICE working paper. Available: www. riceinstitute. org.

Taylor-Robinson, D. C., N. Maayan, K. Soares-Weiser, S. Donegan, and P. Garner (2012). Deworming drugs for soil-transmitted intestinal worms in children: Effects on nutritional indicators, haemoglobin and school performance. *Cochrane Database Syst Rev 7*.

Thomas, D. and E. Frankenberg (2002). Health, nutrition and prosperity: a microeconomic perspective. *Bulletin of the World Health Organization 80*(2), 106–113.

Thomas, D., E. Frankenberg, J. Friedman, J.-P. Habicht, N. Jones, C. McKelvey, G. Pelto, B. Sikoki, J. Smith, C. Sumantri, and W. Suriastini (2004). Causal effect of health on labor market outcomes: Evidence from a random assignment iron supplementation intervention. California Center for Population Research Working Paper.

Tolentino, K. and J. Friedman (2007). An update on anemia in less developed countries. *The American journal of tropical medicine and hygiene 77*(1), 44–51.

UNICEF (2002). Prevention and control of nutritional anaemia: A South Asia priority. *Regional office for South Asia: UNICEF*.

UNICEF & WHO (2012). Progress on drinking water and sanitation: 2012 update.

Van Stuijvenberg, M. E., J. D. Kvalsvig, M. Faber, M. Kruger, D. G. Kenoyer, and A. S. Benadé (1999). Effect of iron-, iodine-, and β-carotene–fortified biscuits on the micronutrient status of primary school children: A randomized controlled trial. *The American journal of clinical nutrition 69*(3), 497–503.

Walker, M. (2003). What is tropical sprue? *Journal of gastroenterology and hepatology 18*(8), 887–890.

Walter, T., I. D. Andraca, P. Chadud, and C. Perales (1989). Iron deficiency anemia: Adverse effects on infant psychomotor development. *Pediatrics 84*(1), 7–17.

WHO, UNICEF, and United Nations University (2001). Iron deficiency anaemia: Assessment, prevention and control. a guide for programme managers. *Geneva: World Health Organization 132*.

World Bank (2013). Data: Land area (sq. km).

Yip, R. and U. Ramakrishnan (2002). Experiences and challenges in developing countries. *The Journal of nutrition 132*(4), 827S–830S.

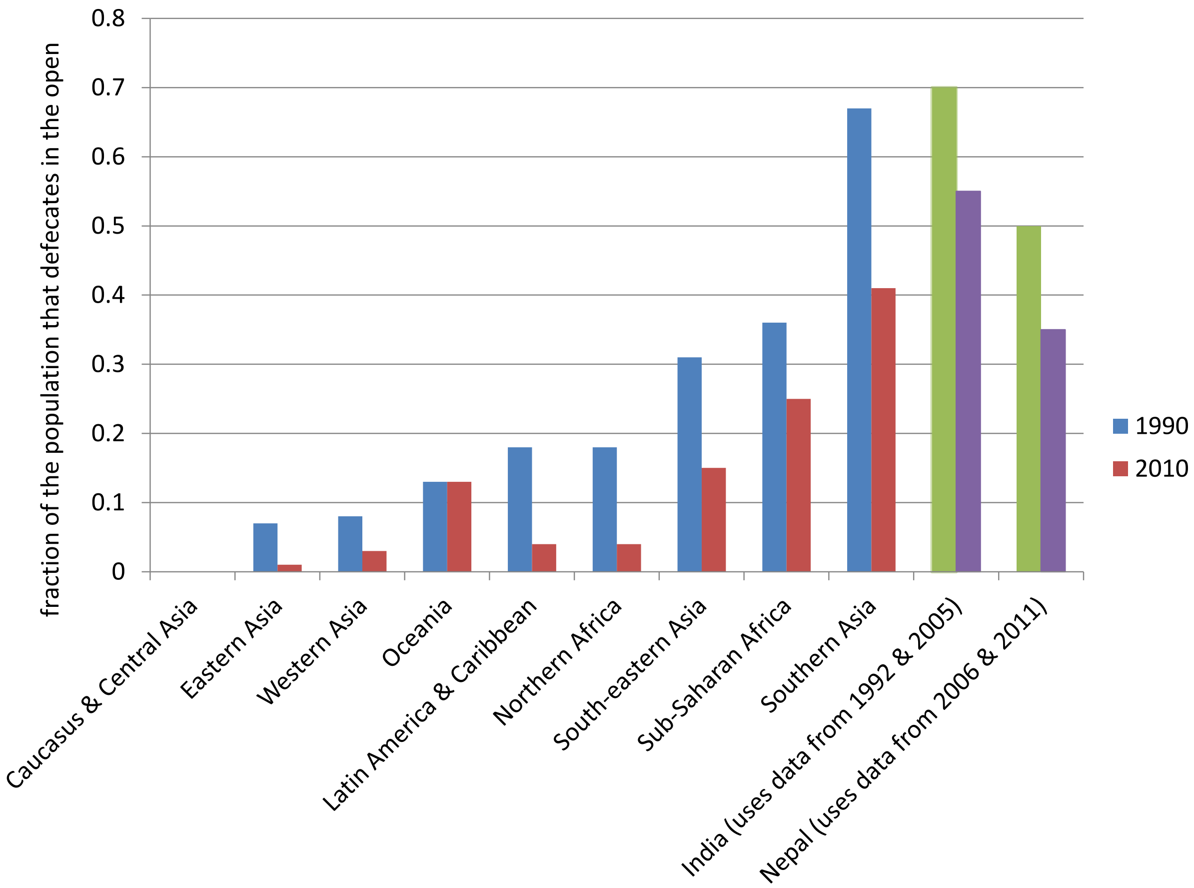


Figure 1: Regional prevalence of open defecation

The 2012 UNICEF & WHO Joint Monitoring Report for "Progress on drinking water and sanitation" is the source of regional estimates of the prevalence of open defecation in 1990 & 2010 (UNICEF & WHO, 2012). For the far right columns, which show household, rather than population level (as in the regional estimates), open defecation in India and Nepal, data are taken from the Demographic and Health Surveys.

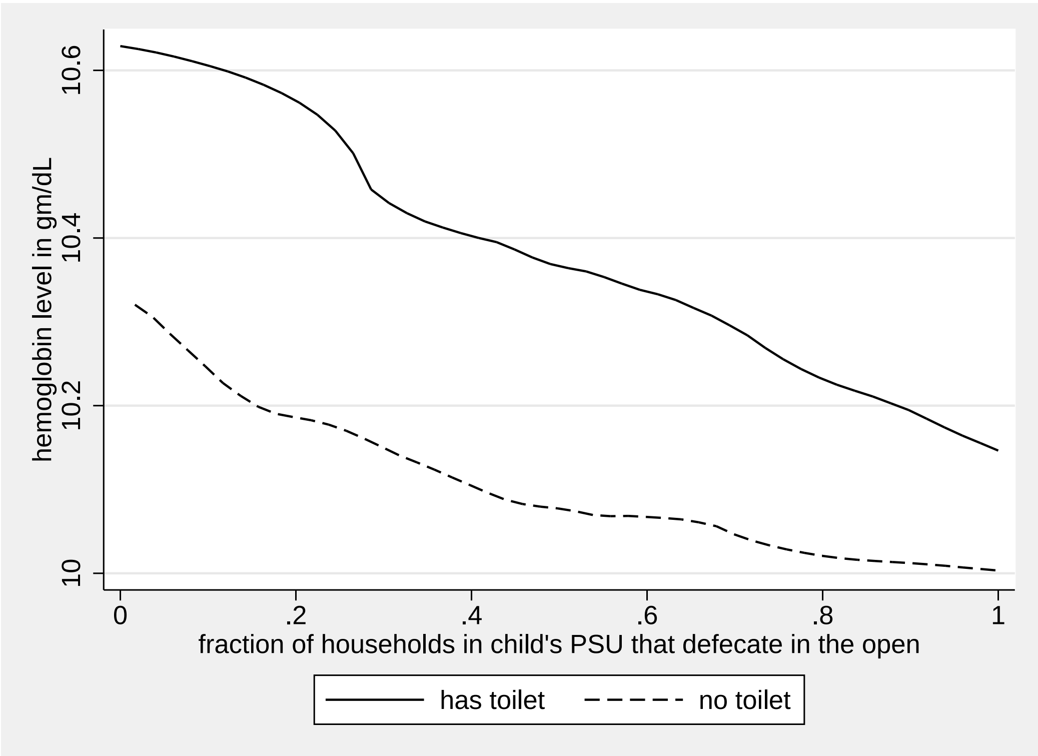
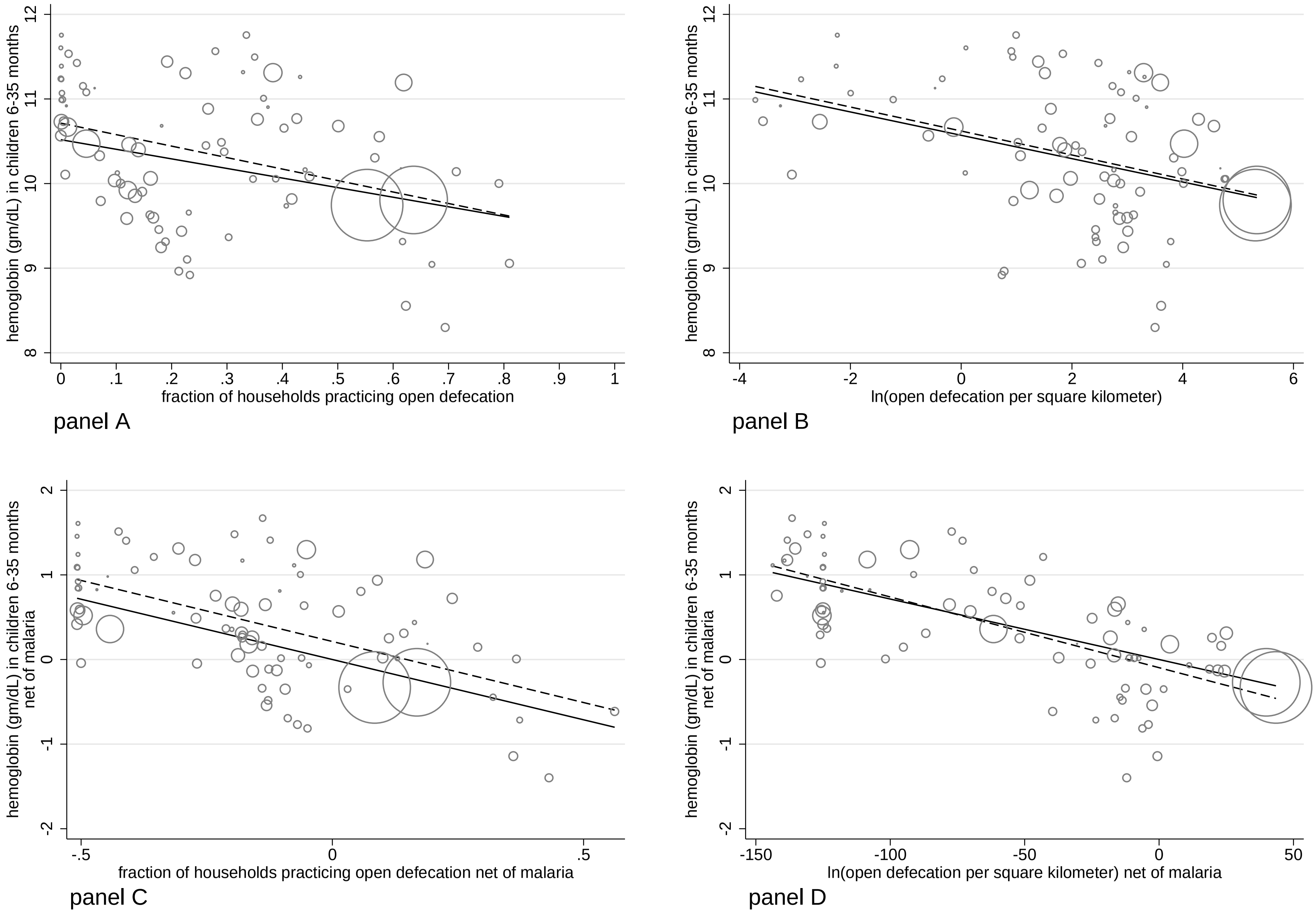


Figure 2: Hemoglobin-community open defecation gradient among Indian children whose households do and do not have latrines

Data are from the India, 2005 DHS. Local polynomial regressions using an epanechnikov kernel.

Figure 3: Cross country relationships between hemoglobin & open defecation



Data are from 81 Demographic and Health Surveys. Solid lines are weighted by the population of the country, dotted lines are not. Scatters are population weighted.

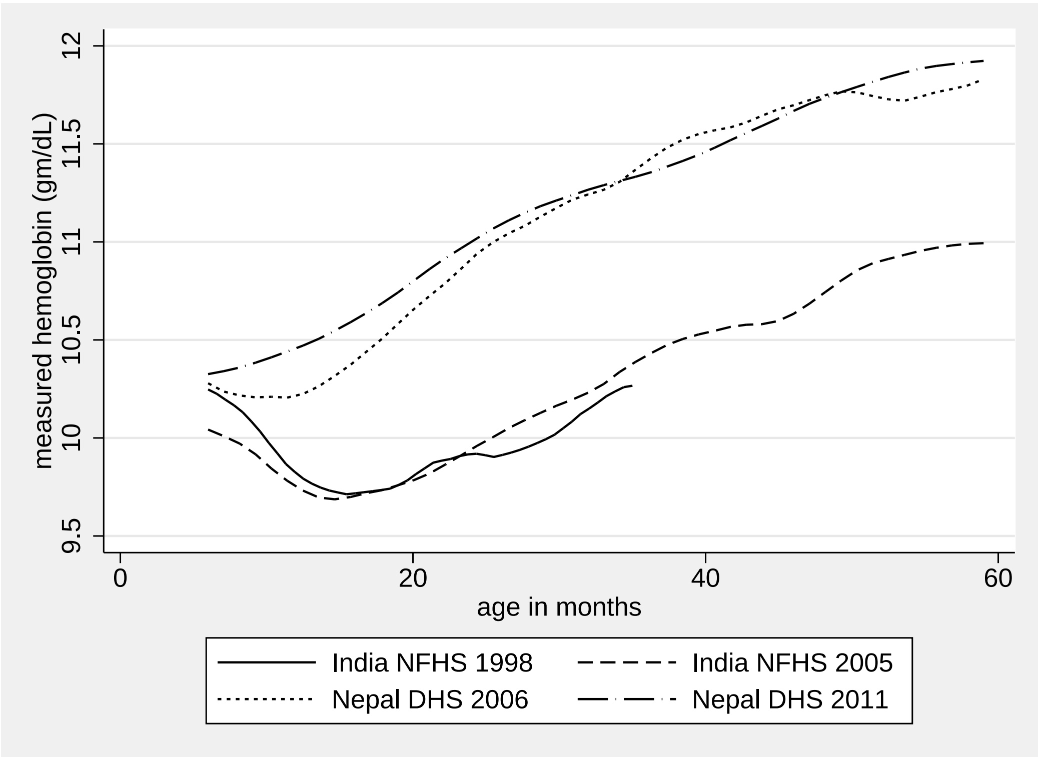


Figure 4: Hemoglobin levels of children between 6 months and 5 years in India and Nepal

Data are from the India 2005 & 1998 DHS and the Nepal 2011 & 2006 DHS. Local polynomial regressions using an epanechnikov kernel.

Table 1: Cross country relationship between hemoglobin and open defecation

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
dependent variable:	average hemoglobin among kids 6 to 35 mo.					average hemoglobin among kids 6 to 59 mo.				
Panel A: uses ln(open defecators per sq km) as measure of open defecation										
ln(open defecators	-0.138***	-0.173***	-0.224***	-0.174*	-0.148***	-0.134***	-0.177**	-0.245***	-0.159*	-0.151***
per sq km)[a]	(0.0265)	(0.0393)	(0.0452)	(0.0676)	(0.0388)	(0.0319)	(0.0505)	(0.0473)	(0.0720)	(0.0406)
estimated malaria		-1.650***	-2.667***	-5.282***	-3.201***		-1.936***	-3.354***	-5.225***	-3.369***
incidence rate[b]		(0.431)	(0.668)	(0.870)	(0.649)		(0.485)	(0.750)	(0.924)	(0.690)
GDP per capita[c]			-0.0189+	-0.0101+	-0.00995			-0.0238*	-0.0120	-0.0125
			(0.0100)	(0.00549)	(0.00814)			(0.0106)	(0.00784)	(0.00823)
region fixed effects				✓					✓	
year fixed effects					✓					✓
n	81	81	81	81	81	77	77	77	77	77
R^2	0.263	0.430	0.518	0.776	0.750	0.194	0.424	0.556	0.767	0.760
Panel B: uses fraction of households that defecate in the open										
fraction of households	-1.133***	-1.425***	-1.533***	-0.728***	-0.790+	-1.065**	-1.343**	-1.430**	-0.699**	-0.705+
openly defecating	(0.227)	(0.258)	(0.327)	(0.176)	(0.418)	(0.366)	(0.463)	(0.476)	(0.217)	(0.406)
estimated malaria		-1.599***	-1.960**	-5.208***	-2.650***		-1.731***	-2.147**	-5.196***	-2.777***
incidence rate		(0.414)	(0.633)	(0.791)	(0.677)		(0.470)	(0.787)	(0.841)	(0.702)
GDP per capita			-0.00732	-0.00714	0.000578			-0.00792	-0.00771	-0.000847
			(0.0104)	(0.00697)	(0.00655)			(0.0120)	(0.00981)	(0.00667)
region fixed effects				✓					✓	
year fixed effects					✓					✓
n	81	81	81	81	81	77	77	77	77	77
R^2	0.233	0.391	0.407	0.783	0.709	0.152	0.340	0.359	0.772	0.718

Observations are DHS surveys in which hemoglobin was measured. Standard errors clustered at the country level (number of countries = 45) and shown in parentheses. Observations are weighted by population in the country year. † $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

[a] ln(open defecators per sq km) is an estimate of the log of the number of open defecators per square kilometer derived by multiplying the fraction of households openly defecating by the population in that country in that year, dividing by the land area in kilometers squared and taking the log.

[b] Malaria incidence rates are taken from Korenromp (2005)'s WHO estimates malaria incidence rates for each country in 2004.

[c] GDP per capita is taken from Heston et al. (2012) which defines the variable as PPP converted GDP per capita (Laspeyres) in hundreds of US dollars, derived from growth rates of c, g, i, at 2005 constant prices.

Table 2: Child level summary statistics in India and Nepal

	India, DHS 2005			Nepal, DHS 2011		
	mean	sd	n	mean	sd	n
hemoglobin level						
age 6-11 months	9.96	1.39	3692	10.32	1.22	226
age 12-23 months	9.93	1.43	4314	10.27	1.3	262
age 24-59 months	9.82	1.51	12226	10.53	1.29	689
breast feeding						
age 6-11 months	0.94	0.24	3687	0.98	0.15	226
age 12-23 months	0.93	0.25	4309	0.96	0.2	262
number of times fed solid or semi-solid food in 24 hours						
age 6-11 months	1.85	1.55	3621	2.78	1.69	226
age 12-23 months	1.95	1.56	4229	2.88	1.67	262
age 24-59 months	2.41	1.59	11999	3.28	1.58	666
fed meat in last 24 hours	0.06	0.24	26750	0.17	0.37	1030
fed dark green leafy vegetables in last 24 hours	0.29	0.45	26790	0.26	0.43	1030
taking iron supplements	0.06	0.23	35650	0.03	0.17	2088
took parasite medicine in last six months	0.15	0.35	35663	0.77	0.42	2082
mother's years of education	5.01	5.07	35851	3.55	4.02	2088
child's household has a toilet	0.53	0.5	33804	0.54	0.5	1995
child's household has electricity	0.71	0.45	33830	0.66	0.47	1995

Table 3: Place specific summary statistics in India and Nepal

	India, DHS 2005			Nepal, DHS 2006			Nepal, DHS 2011		
	mean	sd	n	mean	sd	n	mean	sd	n
PSU level variables									
open defecation, rural PSUs	0.61	0.36	2201	0.59	0.33	178	0.4	0.32	194
open defecation, urban PSUs	0.11	0.21	1649	0.26	0.3	82	0.12	0.18	95
electrification, rural PSUs	0.66	0.33	2201	0.38	0.36	178	0.7	0.33	194
electrification, urban PSUs	0.95	0.11	1649	0.82	0.23	82	0.96	0.08	95
region level variables									
open defecation, rural regions				0.56	0.17	13	0.42	0.13	13
open defecation, urban regions				0.27	0.13	12	0.14	0.09	12
electrification, rural regions				0.39	0.15	13	0.69	0.16	13
electrification, urban regions				0.79	0.17	12	0.95	0.05	12

Table 4: Open defecation predicts hemoglobin level (gm/dL) in India and Nepal

	India, DHS 2005			Nepal, DHS 2011		
	(1)	(2)	(3)	(1)	(2)	(3)
Panel A: Medicine & food controls						
open defecation, PSU level	-0.664***	-0.628***	-0.610***	-0.482***	-0.480***	-0.678***
	(0.0302)	(0.0303)	(0.0335)	(0.121)	(0.122)	(0.143)
taking iron supplements		0.113**	0.0960*		-0.00913	0.234
		(0.0386)	(0.0451)		(0.150)	(0.153)
took parasite medicine in last six months		0.231***	0.214***		0.132+	0.112
		(0.0252)	(0.0306)		(0.0769)	(0.104)
breast feeding			0.0664*			0.530**
			(0.0287)			(0.187)
number of times fed solid or semi-solid food in last 24 hours			0.0440***			-0.0207
			(0.00726)			(0.0346)
ate meat in last 24 hours			0.200***			-0.0656
			(0.0429)			(0.132)
ate dark green leafy vegetables in last 24 hours			0.125***			-0.0430
			(0.0243)			(0.101)
age-in-months×sex dummies	✓	✓	✓	✓	✓	✓
n	35851	35536	26017	2088	2082	1026
Panel B: SES, med., & food controls						
open defecation, PSU level	-0.535***	-0.278***	-0.241***	-0.707***	-0.326+	-0.705**
	(0.0427)	(0.0525)	(0.0594)	(0.154)	(0.174)	(0.216)
electrification, PSU level	-0.0628	0.0134	0.0370	-0.380*	-0.314+	-0.429+
	(0.0487)	(0.0552)	(0.0616)	(0.171)	(0.163)	(0.227)
urban	-0.275***	-0.129***	-0.0937*	-0.198*	-0.123	-0.279*
	(0.0328)	(0.0334)	(0.0370)	(0.0947)	(0.0923)	(0.140)
wealth percentile[a]	0.00442***			-0.00308+		
	(0.000558)			(0.00164)		
mother's education	0.0305***			0.0378***		
	(0.00250)			(0.00970)		
household toilet		0.108***	0.112**			
		(0.0310)	(0.0369)			
mother's education indicators		✓	✓		✓	✓
economic status indicators[b]		✓	✓		✓	✓
medicine & food controls[c]			✓			✓
age-in-months × sex dummies	✓	✓	✓	✓	✓	✓
n	35851	35815	25991	2088	2088	1026

Observations are children aged six months to five years. Standard errors clustered at the PSU level. † $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. [a] The wealth percentile indicates in what percentile of the wealth index, constructed by the Demographic and Health Surveys, a child's household falls among households which have children whose hemoglobin was measured. [b] Economic status indicators include dummy variables for whether or not the child's household owns a radio, a TV, a fridge, a bicycle, a motorcycle, a car, and whether or not it has electricity. It also includes indicators for the type of drinking water the child's house uses (17 options), what type of floor material the child's house has (12 options), as well as what kind of wall (15 options) and roof material (14 options) the house has. [c] Medicine and food controls include whether a child takes iron supplements, took parasite medicine in the six months before the survey, is breast feeding, ate meat in the last 24 hours, ate dark green leafy vegetables in the the last 24 hours, and the number of meals she ate in the last 24 hours.

Table 5: Fixed effects regressions of hemoglobin and open defecation in Nepal's 2006 and 2011 Demographic & Health Surveys

dependent variable: hemoglobin (gm/dL)	(1)	(2)	(3)	(4)	(5)	(6)
open defecation, regional[a]	-1.503**	-1.405**	-1.405**			
	(0.496)	(0.450)	(0.449)			
electrification, regional[a]			-0.000854			
			(0.321)			
open defecation, PSU level				-0.507***	-0.299**	-0.323**
				(0.0865)	(0.0969)	(0.108)
electrification, PSU level						-0.0785
						(0.116)
year 2011	-0.235*	-0.240**	-0.239+	-0.0870+	-0.0537	-0.0443
	(0.0919)	(0.0885)	(0.122)	(0.0514)	(0.0550)	(0.0573)
fixed effects for urban & rural parts of each region	✓	✓	✓	✓	✓	✓
mother's education indicators		✓	✓		✓	✓
economic status indicators[b]		✓	✓		✓	✓
age-in-months × sex dummies	✓	✓	✓	✓	✓	✓
n	6780	6777	6777	6780	6777	6777

Observations are children aged six months to five years. Children from both the Nepal 2006 DHS and the Nepal 2011 DHS are included. Standard errors clustered at the PSU level. † $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. [a] Regional open defecation is the weighted fraction of households in each of 13 regions, separately for their rural and urban parts. The same procedure is used to calculate regional electrification. [b] Economic status indicators include dummy variables for whether or not the child's household owns a radio, a TV, a fridge, a bicycle, a motorcycle, a car, and whether or not it has electricity. It also includes indicators for the type of drinking water the child's house uses (17 options), what type of floor material the child's house has (12 options), as well as what kind of wall (15 options) and roof material (14 options) the house has.

Table 6: Interaction of parasite medicine interaction with open defecation in India

	(1)	(2)	(3)
open defecation × parasite medicine	-0.171**	-0.196*	-0.142†
	(0.0654)	(0.0763)	(0.0742)
open defecation, PSU level	-0.612***	-0.588***	-0.223***
	(0.0315)	(0.0347)	(0.0607)
took parasite medicine in last six months	0.303***	0.289***	0.182***
	(0.0331)	(0.0413)	(0.0400)
own toilet			0.112***
			(0.0369)
food controls[a]		✓	✓
socioeconomic & education controls[b]			✓
age-in-months × sex indicators	✓	✓	✓
n	35663	26017	25991

Data are from the India, 2005 DHS. Observations are children aged six months to five years. Standard errors clustered at the PSU level. † $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. [a] Medicine and food controls include whether a child takes iron supplements, is breast feeding, ate meat in the last 24 hours, ate dark green leafy vegetables in the the last 24 hours, and the number of meals she ate in the last 24 hours. [b] Socioeconomic status indicators include dummy variables for each year of mother's education, dummy variables for whether or not the child's household owns a radio, a TV, a fridge, a bicycle, a motorcycle, a car, and whether or not it has electricity. It also includes indicators for the type of drinking water the child's house uses (17 options), what type of floor material the child's house has (12 options), as well as what kind of wall (15 options) and roof material (14 options) the house has.

Appendix: Demographic and Health Surveys included in the cross country analysis

sub Saharan Africa	
Angola	2006
Angola	2011
Benin	2001
Benin	2006
Burkina Faso	2003
Burkina Faso	2010
Burundi	2010
Cameroon	2004
Cameroon	2011
Congo (Brazzaville)	2005
DRC	2007
Ethiopia	1997
Ethiopia	2003
Ghana	2003
Ghana	2008
Guinea	2005
Lesotho	2004
Lesotho	2009
Liberia	2008
Liberia	2011
Madagascar	2003
Madagascar	2008
Madagascar	2011
Malawi	2004
Malawi	2010
Malawi	2012
Mali	2001
Mali	2006
Mozambique	2011
Niger	2006
Rwanda	2005
Rwanda	2007
Rwanda	2010
Sao Tome and Principe	2008
Senegal	2005
Senegal	2008
Senegal	2010
Sierra Leone	2008
Swaziland	2006
Tanzania	2004
Tanzania	2007
Tanzania	2009
Tanzania	2011
Uganda	2000
Uganda	2006
Uganda	2009
Zimbabwe	2005
Zimbabwe	2010
East Asia	
Cambodia	2000
Cambodia	2005
Cambodia	2010
Timor-Leste	2009
Europe	
Albania	2008
Armenia	2000
Armenia	2005
Moldova	2005
Ukraine	2011
Latin America	
Bolivia	1998
Bolivia	2003
Bolivia	2008
Guyana	2009
Haiti	2000
Haiti	2005
Honduras	2005
Peru	1996
Peru	2000
Peru	2003
North Africa	
Egypt	2005
South Asia	
Bangladesh	2011
India	1998
India	2005
Nepal	2006
Nepal	2011
West Asia	
Azerbaijan	2006
Jordan	2002
Jordan	2007
Jordan	2009
Kazakhstan	1995
Kazakhstan	1999
Kyrgyz Republic	1997
Uzbekistan	1996