The impact of piped water provision on infant mortality in Brazil: A quantile panel data approach

Shanti Gamper-Rabindran

Shakeeb Khan

Christopher Timmins[1]

June 27, 2007

Abstract

We examine the impact of piped water on the under-1 infant mortality rate (IMR) in Brazil using a novel econometric procedure for the estimation of quantile treatment effects with panel data. The provision of piped water in Brazil is highly correlated with other observable and unobservable determinants of IMR – the latter leading to an important source of bias. Instruments for piped water provision are not readily available, and fixed effects to control for time invariant correlated unobservables are invalid in simple quantile regression framework. Using the quantile panel data procedure in Chen and Khan (2007), our estimates indicate that the provision of piped water reduces infant mortality by significantly more at the higher conditional quantiles of the IMR than at the lower conditional quantiles (except for cases of extreme underdevelopment). These results imply that targeting piped water intervention in areas with higher conditional quantiles of the IMR, when accompanied by a basic level of other public health inputs, can achieve significantly greater reductions in infant mortality.

JEL Codes: I18, H41, Q53, Q56, Q58

Keywords: Infant mortality, piped water supply, quantile fixed effects, heterogenous program impact, distribution of public goods

---

[1] Assistant Professor, University of Pittsburgh, and Associate Professors, Duke University, respectively. Corresponding author: Shanti Gamper-Rabindran. Email: shanti@gspia.pitt.edu. Phone: 412-648-8266 Fax 412-648-2605. Acknowledgments: We thank colleagues at IPEA, John Briscoe, Daniel Kammen, Grant Miller, Kara Nelson, Narayan Sastry, Werner Troesken, seminar participants at Yale School of Forestry and Environmental Studies, University of Pittsburgh and the Population Association of America annual meeting. Errors are ours.

1 Introduction

The Millennium Development Goals aim to reduce by two-thirds the under-five child mortality rate by 2015 from the base year 1990 (United Nations, 2005). Diarrhea caused about 22% of the under-five deaths in 2000 (Black et al, 2003).[2] About 1.5 million child deaths (or 88% of those from diarrhea) are caused by ingestion of unsafe water, inadequate availability of water for hygiene, and lack of access to sanitation (Black et al, 2003). A proposed strategy to achieve the Millennium Development Goals of reducing child mortality is to improve access to safe drinking water. This raises an important policy question – can the provision of piped water from the network, hereafter "piped water", reduce the infant mortality rate?[3] For those populations at greatest risk (i.e. in areas that suffer severe infant mortality rates) can this provision reduce infant mortality rates, or is the provision of piped water effective only when accompanied by complementary income-related inputs at the household or community level?

In situations involving extreme inequality, it is possible for simple conditional mean estimates to mask the answer to these questions. Quantile estimation, which allows us to recover the marginal impact of piped water on various quantiles of the conditional distribution of the IMR, can address this problem. Quantile regression is, however, not easily adaptable to dealing with problems of endogenous regressors. This presents a difficulty for most policy analyses, since policies are seldom applied randomly. When valid instruments are available, endogeneity can be addressed with instrumental variable quantile techniques (Abadie et al., 2002; Arias et al., 2001; Chernozhukov and Hansen, 2005). The practical problem is that there are often no good instruments for many policies. The usual statistical approach in mean regression is to exploit panel variation and estimate fixed effects to control for time invariant sources of

[2] These figures are for the 42 countries with 90% of the worldwide under-5 deaths in 2000 (Black et al., 2003).

[3] Our study is limited to the following measure: percentage of households that receive piped water from the network. We do not have information on the quality of that piped water. Furthermore, we do not evaluate the effectiveness of piped water interventions relative to other water-related interventions. Mintz et al. (2001) argue that "decentralized approaches to making drinking water safe, including point-of-use chemical and solar disinfection, safe water storage, and behavioral change merit far greater priority for rapid implementation."

correlated errors. However, this approach is not applicable using standard quantile techniques.[4]

Using a new approach to quantile regression with panel data developed by Chen and Khan (2007), we examine the impact of provision of piped water on the under-1 infant mortality rate (IMR) at various quantiles of the conditional IMR distribution using panel data for 3568 census units in all Brazil. We describe the effect of the treatment on various quantiles of the outcome distribution, making no assumption about the joint distribution of the treated and untreated distributions. Our interpretation follows that of Abrevaya (2001) and Bitler et al., (2005).

We find that an increase of one percentage point in the number of households receiving piped water in the group of counties with poor development indicators in the period 1980-1991 causes a decline of 0.86 deaths per 1,000 live births at the 90th percentile of the conditional IMR, but a decline of only 0.37 deaths at the 10th percentile.[5] The marginal effect at the mean (i.e., 0.68 deaths per 1,000 live births) turns out to provide a poor indication of the effect of water on much of the IMR distribution. The most important implication of this result is that the impact of a piped water provision policy is determined in large part by how those piped water connections are distributed. There is tremendous payoff to targeting water provision to the areas with the highest IMR (both conditional and unconditional). In practice, however, piped water interventions have tended to be places with good indicators of development and which are low in the conditional IMR distribution.

Our paper makes two methodological contributions to the program evaluation literature in developing countries. First, by using novel quantile techniques, we examine whether the provision of piped water can reduce infant mortality at the higher tails of the conditional quantile distribution. A priori, it is unclear whether the provision of piped water, without sufficient complementary health inputs, will yield a reduction IMR at these quantiles. Previous studies'

---

[4] Differenced regression cannot be applied in the quantile regression context, and simple fixed effects suffer from incidental parameters bias unless the panel is very long in the time dimension.
[5] These indicators are described in Table 4.

focus on the impact at the mean of the conditional distribution may obscure this policy-relevant heterogeneity.

Second, by applying panel data techniques to quantile estimation, we can estimate the impact of piped water on IMR while controlling for potential time invariant confounders. Areas with fewer piped water connections are also high IMR areas. These areas may suffer from systematic underreporting of infant deaths (Victora and Barros, 2001). At the same time, areas with more piped water connections are likely to benefit from other superior health inputs (unobservables in our study, like access to variety of types of medical care, nutritional supplements, public health infrastructure, etc…) (Jalan and Ravallion, 2003; Weinreb, 2001). Our estimates will not suffer from the downward bias arising from the systematic underreporting of deaths or the upward bias arising from these time invariant inputs.[6] At present, only a few, albeit important papers, have applied quantile regression to program evaluation in developing countries (Djebbari and Smith, 2005) and fewer still have applied strategies to address time invariant confounders with the context of quantile regressions.

While instrumental variables may not always be available, the proliferation of quality panel data means that our methodological approach can be widely applied to the evaluation of other programs that provide health inputs or other public goods in developing countries. Our task of evaluating the impact of piped water on IMR shares two key characteristics with the evaluation of programs that provide health inputs in developing countries, such as the provision of nutritional supplements or medical assistance to populations at risk. First, from the policy perspective, it is important to understand the impact of these programs on the subpopulations that are most at risk; if unobservables are important determinants of the outcome variable, these subpopulations will tend to occupy the tails of the conditional outcome distribution. Mean impacts will fail to capture heterogeneous impacts across the conditional distribution. Second, the evaluation of these programs is complicated by their systematic placement of provision in areas that

[6] The quantile panel data technique we employ, like other fixed effect models, cannot correct the bias arising from time varying unobservables.

receive superior health inputs. If these inputs are unobserved by the econometrician, they will cause an upward bias in the measurement of positive program impacts. At the same time, the persistent underreporting of outcome variables (e.g., mortality in higher mortality areas) (Victora and Barros, 2001), may attenuate the relationship between health inputs and mortality. Random assignment of treatment and plausible instruments are often not available to assist program evaluation in developing countries. With the increasing availability of panel data in these countries, however, the panel data approach provides a promising strategy to address the issue of bias arising from unobservables (albeit only time invariant ones) within the context of quantile regressions.

2 Piped water and infant mortality in Brazil

Brazil serves as a case study for the impact of piped water on infant mortality for three reasons. First, diarrheal diseases are an important cause of infant mortality, accounting for 8% of infant death in Brazil in 1995-7 (Victora, 2001). In Northeast Brazil, the poorest area in the country, diarrhea accounted for 15% of infant mortality (Victora, 2001). Second, under-1 infants in Brazil are susceptible to water-borne diseases due to the relatively short duration of breastfeeding. Diarrhea is likely to increase when the infant is first exposed to supplemental liquids or solids (Sastry and Burgard, 2002), usually at ages below 1 year old. (Sastry and Burgard, 2002). The 1989 Brazilian National Health and Nutrition Survey indicates that only 29.5% of infants aged 0-5 months were exclusively breastfed and 36.3% of those aged 0-23 months were breastfed (Senauer and Kassouf, 2000). In 1996, the Brazil-wide estimate of the duration for breastfeeding (both exclusive and supplemental) was 8.2 months (Sastry and Burgard, 2002).[7] Third, Brazil (particularly its high infant mortality areas in the Northeast) shares several characteristics with other developing countries that make our results potentially transferrable. These characteristics include its high

[7] The average duration of breastfeeding did not differ dramatically between the Northeast and the rest of Brazil (Sastry and Burgard, 2002).

IMR, poor provision of piped water, low literacy and income levels, and its location in tropical and subtropical climate zones.

Piped water supply can have direct and indirect influences on infant mortality. Infants in Brazil are likely to be exposed to contaminated water and poor hygiene from prepared food from an early age (Merrick, 1985) due to the short duration of breastfeeding (Anderson, 1981 cited in Merrick, 1985; Sastry and Burgard, 2002). Piped water supply can also reduce infant mortality indirectly when caregivers are able to devote more time to childcare instead of water collection activities.

2.1 Marginal effects of piped water – mean versus quantiles

We use quantile techniques to recover the marginal impact of piped water on various parts of the conditional IMR distribution. In contrast, previous studies on piped water focus on the conditional mean of that distribution (Sastry, 1996; Merrick, 1985; Jalan and Ravallion, 2003). Only under the assumption that the marginal effect of piped water is a simple "common effect" or "location shift" will the impact at the mean be the same as the impact for the entire distribution (Heckman et al., 1997, Abadie et al., 2002). In other words, under the “common effect” assumption, the piped water intervention has the same impact on everyone with the same observed characteristics (Heckman et al., 1997).

Papers on health inputs have shown that estimates at the mean may obscure heterogeneous impacts at the various quantiles of the conditional distribution. Moreover, the heterogeneity in the conditional distribution of the outcome variable is relevant for public policy. For example, Abrevaya (2001) finds that prenatal care in the US has a significantly higher impact at lower quantiles of the conditional distribution of infant birthweight than at the higher quantiles. Moreover, he finds that the black-white differential in birthweight is larger at the lower conditional quantiles of birthweight.

Heterogeneity in the impact of piped water is relevant for policy decision of piped water placement. On the one hand, targeting piped water to vulnerable households may improve their welfare significantly. Households or communities with low income typically have the fewest public resources for children's health.[8] In such cases, we would expect piped water to have greater protective effect among households or communities with lower incomes. On the other hand, targeting piped water to vulnerable households may be necessary but not sufficient to improve their welfare. In particular, their limited income or education may constrain their ability to exploit the benefits from piped water supply. In that case, water supply placement would need to be accompanied by other interventions (Jalan and Ravallion, 2003).

In exploring the impact of piped water on IMR, our study explores two types of heterogeneity which call for distinct policy responses: (1) heterogeneity along observable dimensions such as income and (2) heterogeneity due to unobserved variables. The policy response for the first type of heterogeneity is to target along observables such as income and education. The policy response to the second type of heterogeneity is more challenging. It would not be sufficient to simply consider income, education, and sewage in defining "vulnerable populations". Instead, in their task of allocating water, policy-makers need to look for other factors (i.e., unobserved factors in our analysis) that make IMR high.

The first type of heterogeneity can be explored by standard techniques in the literature i.e., by allowing the marginal impact of water to vary by the income variable. However, a mean regression with interaction terms would ignore the second type of heterogeneity. In contrast, the quantile techniques allow us to explore the second type of heterogeneity. A priori, it is unclear whether the marginal effect of water is greater in higher or lower income communities. Similarly, a priori, it is unclear whether, controlling for observables such as income, education and sewage network, the marginal effect of water is greater at higher or lower percentiles of the conditional distribution of the IMR.

Previous studies suggest a complex relationship between health status, water supply and socionomic status.[9] Shuval et al. (1981) propose a four-stage

---

[8] Thomas and Strauss (1992) make this argument for maternal education.

[9] Consider, for example, hygiene behavior, which our study ignores but which can influence whether the provision of water supply translates to health benefits. It is likely that the provision of

threshold-saturation model to explain the relationship between health status, water supply and socioeconomic levels reported in several empirical studies with seemingly contradictory results.[10] Shuval et al. (1981) propose that at the first stage, i.e., below a threshold of socioeconomic development, the provision of water does little to improve the health status of the community. Individuals have low disease resistance due to their extremely poor nutrition and personal hygiene and their exposure to multiple and simultaneous routes of disease transmissions. The provision of water alone, which addresses only one route of disease transmission, does not have a strong impact on health. Shuval et al.'s (1981) argument echoes that of Briscoe (1984) – i.e., the improvements in drinking water supply in Matlab, Bangladesh did not cause major reductions in cholera incidence because complementary interventions were not undertaken to eliminate other important, albeit secondary, routes of cholera transmission (e.g., the ingestion of polluted water during bathing). Similarly, Esrey et al. (1992) find that water supply had a significant health impact only when accompanied by the presence of latrines in their study of infants in Lesotho.[11]

At the second stage, above that threshold, but below the saturation point, socioeconomic development improves the standard of living and reduces the exposure to infection. (Shuval et al., 1981) At this level of socioeconomic development, communities have a strong health response to investments in water supply. At the third stage, as communities develop further, they move towards a saturation point, whereby improvements in water supply have only a

---

water supply encourages the adoption of hygienic behaviors. Cairncross (1990) argues that the provision of water leads to health impacts only when accompanied by the adoption of hygienic behavior. Citing Esrey et al. (1985), . Cairncross (2003) argues that handwashing thus turns out to have an even greater impact on diarrheal disease than water supply or sanitation. Nevertheless, as noted by Cairncross (2003), "a convenient water supply makes handwashing easier to practice and hence more likely. Indeed, it has been confirmed by observation in developing countries that mothers of young children are more likely to wash their hands at critical moments if they have a piped water supply (Curtis et al., 1995)."

[10] Shuval et al. (1981) use country-level data for 65 developing countries from 1962. Life expectancy at birth measures health status, adult literacy rate measures socioeconomic status, and the proportion of the urban population having access to water supply by either household tap or standpipe measures the sanitation level. Shuval et al. (1981) reports that these data, though imperfect, are consistent with their model.

[11] Esrey et al. (1992) examine 119 infants who lived in 20 villages in Lesotho from a 6 month period in 1984-1985.

small impact on health. At the fourth stage, beyond the saturation point, communities have reached high levels of socioeconomic development. Improvements to water supply would not cause further improvements in health status. (Shuval et al., 1981)

Previous studies show diverging results on the interaction between piped water and income. Whether piped water serves as a complement or substitute to household and community inputs may be specific to the level of income and education and overall institutional environment. In their study of 33,000 rural households in India in 1993-1994, Jalan and Ravallion (2003) find that while piped water did cause an overall reduction in diarrheal incidence, households in the bottom 40 percent of the income distribution did not experience significant health gains.[12] In their study of Brazil in 1974/5, Thomas and Strauss (1992) find that children in high income urban households benefit more from the availability of sewerage services and electricity. In contrast, several studies report that households' input and public infrastructure serve as substitutes. Thomas et al. (1991) find that children of uneducated mothers gained most from sewage networks in Northeast Brazil. Barrera (1990a) finds that children of less educated mothers in the Bicol region of the Phillipines benefit more from water connections and the absence of excreta in the environment.[13]

2.2 Selective placement of piped water

Studies relating water supply and health that fail to control for the selective placement of water supply would likely overstate the protective effect of water (Zwane and Kremer, 2007). Piped water is likely to be placed in areas that enjoy superior medical care provision, and where higher incomes are used to purchase other health-related inputs (Jalan and Ravallion, 2003; Weinreb, 2001). Both represent factors that contribute to low IMR (Rosenzweig and Wolpin, 1986). Our

[12] Jalan and Ravallion (2003) argue that "policymakers trying to reach children of poor families – who are typically the most prone to disease – will need to do more than making facility placement pro-poor. The incidence of health gains need not favor children from poor families even when the placement favors the poor."

data indicate that piped water in Brazil is systematically placed in areas with superior observables. The correlation between the water and income variables is 0.71, 0.73, 0.78, and 0.61 and that between the water and education variables is 0.60, 0.63, 0.71, and 0.58, in 1970, 1980, 1991 and 2000, respectively.[14] It is therefore likely that the placement of piped water is also correlated with unobservable determinants of infant health.

To overcome the estimation problem of selective program placement, studies from developed countries exploited the exogenous timing of water-related interventions to identify their impact. Troesken (2001) finds that municipal water provision in American cities around the early 20th century reduced typhoid rates in blacks. Cutler and Miller (2005) report that chlorination interventions in 19 US cities reduced infant mortality. Watson (2006) finds that a ten percentage point increase in the fraction of homes in American Indian reservations with sanitation improvements reduced infant mortality by 0.51 deaths per 1000 births.
However, few studies from developing countries (even those that focus on mean results) have been able to correct for non-random program placement. In their study of Bangladeshi and Filipino villages, Lee et al. (1997) correct for the selection bias stemming from conditioning on surviving children, but take the placement of piped water as given. In their Brazilian studies, Sastry (1996) and Merrick (1985) report positive association between piped water supply and infant mortality, but are not able to address the issue of program placement.[15] The study by Jalan and Ravallion (2003) uses propensity score matching techniques as a strategy to correct for the selective placement of piped water among rural Indian households in 1991. Comparing households with and without piped water, but which are similar on observable dimensions (and, by assumption, on

---

13 In the same study, Barrera (1990a) finds that children of more educated mothers derive greater benefits from health care facilities and toilet connections. Thomas et al. (1991) find that children of uneducated mothers gained least from health care facilities.

14 These variables are defined in section 5.1.

15 Merrick (1985) uses 1976 cross-sectional in Brazil data to estimate a structural model relating infant mortality to factors such as household-level access to piped water, state-level piped water supply, maternal and paternal education, and income. Merrick (1985) obtained piped water supply data from the 1970 Census that divided Brazil into 117 geographical units. In order to match the data to the Pesquisa Nacional Amostra de Domicilios (PNAD) household data, he was forced to aggregate the variable up to 25 observations corresponding to 25 states.

unobservable dimensions), they find that the incidence of diarrheal diseases is higher in households without piped water. In contrast, studies testing the impact of point-of-use water treatment have been able to implement randomized trials. Clasen et al. (2004), Conroy et al. (1996), and Crump et al. (2005) find positive health impacts from the use ceramic filters, solar disinfection and chemical disinfectants, respectively. Kremer et al. (2006), using randomized trials to evaluate the impact of protecting naturally occurring spring water, find no significant child health effects but note a limited ability to reject the null hypothesis of no effect because of weak power.

2.3 Measurement error in IMR

Measurement error in the IMR poses a second problem in studies that investigate the relationship between health inputs and infant mortality, albeit the direction of bias is opposite to that discussed above for program placement. In particular, places that suffer from higher infant mortality rates may suffer more severe under-reporting of those rates. Even if there were an underlying negative relationship between the presence of piped water and infant mortality, the underreporting bias may conceal such a relationship. In the case of Brazil, Victora and Barros (2001), citing Simões (1999), note that under-reporting of infant deaths in the Northeast (where provision of piped water is low) is about 66.7%, while the under-reporting in the Southeast (where provision of piped water is high) is only 6.5%.[16]

2.4 Strategies to address non-random program placement and measurement error

To overcome the issue of non-random program placement, quantile studies from developed countries have been able to rely on experimental design such as in the evaluation of welfare reform or job training programs (Bitler et al.,

[16] The under-registration of infant deaths is estimated to be 52.2% in the North, 13.6% in the South, 23.9% in the Center-West, and 43.7% nationally. Most deaths that are not registered occur in the rural areas of the North and Northeast where rates of infectious diseases are higher (Victora and Barros, 2001).

2005 and 2006) or instrumental variables such as in evaluating the impact of childbearing on income[17] (Abadie et al., 2002), the returns to education (Arias et al., 2001), returns to job training programs (Chernozhukov and Hansen, 2005). In contrast, only a few quantile studies from developing countries have been able to rely on experimental design or instrumental variables. Djebbari and Smith (2005) use random assignment experimental data to examine the distributional impact of Mexico's program of education, health and nutrition (PROGRESA). They find that the program had a smaller impact on wealth and nutrition for households in the lower tail of the wealth and nutrition distribution.

A few studies, looking only at developed countries, have begun to explore the use of panel data in the context of quantile regressions. For example, Abrevaya and Dahl (2006) examine the impact of prenatal care and smoking on infant birthweight using panel data on maternally-linked births. They assume a correlated random effects model as is done here and in Chen and Khan (2007), but also impose the further restriction of a linear structure on the individual specific effect. In this and other important policy contexts, randomized placements and instrumental variables are not readily available. In these situations, the use of panel techniques has the potential to correct the estimation bias from selective placement and systematic measurement error.

3 Why means mask quantile results?

In this section we illustrate estimating mean effects can be significantly different from quantile effects. To keep the discussion simple, we focus here on the cross-sectional case. In particular, consider the linearly heteroskedastic model: $y_i = \beta_0 + x_i\beta_1 + x_i\Psi\varepsilon_i$, where y_i measures the infant mortality rate in county i and x_i measures the percentage of households there with access to piped water. We assume for this discussion that ε_i is independent of x_i (although relaxing this assumption is a major focus of the rest of the paper). Let μ_ε denote the mean of ε_i and let ρ_θ denote the θ^{th} quantile of ε_i. The marginal effect associated with the

[17] "Childbearing reduces the lower tail of the income distribution considerably more than other parts of the income distribution." (Abadie et al., 2002).

conditional mean function (which would be estimated were we to use simple OLS) is of the form ($\beta_1+\Psi\mu_\varepsilon$), whereas the marginal effect associated with the θ^{th} quantile is ($\beta_1+\Psi\rho_\theta$). The differences between these two measures will generally depend upon the skewness of the distribution of ε_i.

For example, if Ψ positive and the distribution of ε_i is skewed right, then the marginal effect of x_i associated with the mean will exceed that associated with the median and the lower quantiles. On the other hand, if the distribution is skewed left, the reverse will be true – marginal effects associated with the mean and higher quantiles will exceed the marginal effect attained from OLS.

Brazil is a country characterized by extreme inequality. Its Gini coefficient has risen steadily from 0.56 in 1970 to 0.63 in 1990, making it one of the most unequal nations in the world. Even within-region allocations are unequal – the 1988 Gini coefficient in the Northeast (Brazil's poorest region) was 0.64. Given this unequal distribution of resources, we should not be surprised to find an asymmetric distribution for the unobservable determinants of the infant mortality rate across Brazil. This makes clear the need for the quantile panel data approach we pursue in this paper.

4 Data

We use newly available census data published by the Brazilian Institute for Economic Analysis (IPEA). These data are reported at the level of minimally comparable areas (MCA's) for the years 1970, 1980, 1991 and 2000. Previously, census data were available at the *municipio* or county level, which is the smallest political division in Brazil (Alves and Beluzzo, 2004). Changes in county boundaries between the decades had limited the comparability of the census data. To overcome this limitation, IPEA created the MCA dataset, in which geographical units share common boundaries across the decades. The MCA boundaries correspond to county boundaries for those counties whose borders did not change between 1970 and 2000. For those counties that changed their borders between 1970 and 2000, neighboring counties were dissolved into one larger MCA. Data from households were then aggregated up to the MCA level for 1970, 1980, 1991 and 2000.

The MCA dataset divides Brazil into 3569 MCAs, a number which compares favorably with the 4500 counties in Brazil in 1998 (Mobarak et al., 2004) and 5560 in 2000 (Alves and Beluzzo, 2004).[18] While the MCA dataset is imperfect in that it sometimes aggregates several counties (which may differ in their policy and institutional context), we believe that this dataset represents the best demographic panel dataset currently available for Brazil. The finer resolution of this MCA data relative to other available Brazilian panel census data lessens the degree of within unit heterogeneity.[19]

Table 1 presents the summary statistics. The mean infant mortality rate declined from 125 deaths per 1000 live births in 1970, to 87 deaths in 1980, to 49 deaths in 1991 and to 34 deaths in 2000. At the same time, we see improvements in other development indicators. The percentage of households with piped water has increased fourfold from a mere 15% in 1970 to 62% by 2000. The percentage of households connected to the sewage network, starting from a lower baseline of 5% in 1970, has increased six-fold to 29% by 2000. Total fertility rate has more than halved from 5.9 births in 1970 to 2.8 births by 2000. Both the income-related Human Development Index and the education-related Human Development index show improvement between 1970 and 2000.

5 Method

5.1 Estimation

Our dependent variable is the number of deaths of infants under one year of age per thousand live births. Our analysis focuses on all-cause infant mortality. Brazilian vital statistics data (except when the information is specifically collected by researchers) are notoriously unreliable on cause-specific deaths,

---

[18] In the 1980s, Brazil had 4088 municipalities, with an average population of 29,800 and an average area of 2118 km^2 (Sastry, 1996). We drop one observation in our analysis because of missing values.

[19] Potter et al. (2002) use the previous version of decennial data, terminating in 1991, that divides Brazil into 518 microregions. Another data source, the PNAD, suffers from municipio boundaries that are not consistent from one survey to another.

and the unreliability is worse in high mortality areas (Sastry and Burgard, 2002). By focusing on infant mortality, we avoid the potential bias inherent in studies that examine child health. Studies that use child health (e.g. height for weight scores) need to correct for the selection on surviving children in order to avoid underestimating the overall impact of piped water on child health (Lee et al., 1997).[20] We interpret the coefficient on piped water to capture the impact of piped water on infant mortality, typically through increased risk of death from diarrheal diseases.

Our study is limited to the analysis of one aspect of the quantity of piped water. The definition for the water variable is the percentage of households with piped water from the general network.[21] As in Sastry (1996) we focus on households' source of water and not on the type of connection. Households with water from the network may have water connections through internal or external plumbing. Sastry (1996) reports that infant mortality levels are more strongly correlated with the source of water than the type of water connection.[22] While the absence of information on the quality of water is a limitation in our study, "the few studies that have considered both the quality and quantity of water, find that water quantity has a greater impact than water quality on health and mortality (Sastry, 1996; citing Bourne, 1984, Esrey and Habicht, 1988, and Victora et al, 1988),

In addition to piped water, we include several covariates to account for other time-varying factors that influence the IMR. Lower fertility reduces the infant mortality rate, in part through the positive effect of birth spacing on child mortality (Barnum, 1988).[23] Total fertility rate is a measure that summarizes the

[20] Child health data from the PNAD, 1996 Demographic and Health Survey and 1989 the Brazilian Health and Nutrition fail to provide municipal-level information. Cause-specific vital statistics data are not publicly available for all of Brazil.

[21] The IPEA definition is "numero domicilios com água canalizada rede general".

[22] Our dataset does not have information on households with other sources of water (e.g. well water) and the type of connections in the households (internal or external plumbing). Victora et al. (1988), in their study of two metropolitan areas in Southern Brazil, report that "compared to those with water piped to their house, those without easy access to piped water were found to be 4.8 times more likely to suffer infant death from diarrhea and those with water piped to their plot but not to their house had a 1.5 times greater risk."

[23] Higher fertility influences infant mortality through the effect of birth spacing on child mortality. "A cross-country study finds that children born before a two year birth interval have a 60% higher

rate of childbearing in a year. It is derived by summing the age-specific birth rates for a population of women in a given period. That variable is available at the county-level for 1991 and 2000 only and at the region level for 1980 and 1970.

Maternal education, by improving mother's access to health-related information and her ability to make better use of health inputs, influences the reduction in the infant mortality rate. (Sastry, 1996 citing Barrera (1990a), Rosensweig and Schultz (1982), and Thomas et al (1991)) In the absence of women-specific education or literacy data, we use the education-based Human Development Index (HDI_education) variable. The HDI_education variable has been constructed by IPEA from a 2:1 weighting of the index for literacy rate and the index for school attendance rate.[24] We add as a covariate the income-based Human Development Index (HDI_income), as higher income levels are associated with improved chances for child survival. [25] (Sastry, 1996 citing Merrick 1985, Thomas et al, 1990 and Victora et al., 1986) We also add as a covariate the percentage of households connected to the regular sewage network, since poor sanitation contributes to reductions in infant mortality[26] (Habicht et al., 1988). Finally, our panel data procedure controls for county-specific time invariant characteristics. One such characteristic that influences infant mortality is the climate. Greater seasonality in temperature and precipitation is associated with greater infant mortality from infectious diseases (Sastry, 1996).[27]

The basic panel data model to be estimated is of the form:

---

mortality risk in the neonatal period and 100% in the remainder of the period compared with children born after a two year birth interval (Barnum, 1988; citing Pebley and Millman (1986).

[24] The HDI_education variable includes current schooling, which captures MCA-level investment in education of children. The index of literacy rate or the index of school attendance rate = (observed rate – minimum rate) / (maximum rate – minimum rate).

[25] The definition for HDI_income = ln (observed value of RFPC) – ln (lower limit of RFPC) / [ln (upper limit of RFPC)-ln(lower limit of RFPC) where RFPC is the family per capita income.

[26] Victora et al. (1986) and Victora et al. (1988), both cited in Sastry (1996), find that "household toilet facilities are related very weakly to child mortality risks.

[27] In their examination of the trends in diarrhea prevalence and treatment in Brazil between 1986 and 1996, Sastry and Burgard (1992) report that while treatment with oral rehydration therapy (ORT) increased greatly, there was a very modest decline in diarrhea prevalence in Brazil over this ten year period. The authors conclude that the rise in ORT did not reduce the prevalence of diarrhea. In contrast, Victora et al (1996) report than ORT played a larger role than income, education, and access to water in the sharp decline in infant deaths due to diarrhea in the 1980s.

(1) $$y_{i,t} = \alpha_i + x'_{i,t}\beta + \varepsilon_{i,t} \qquad t = 1, 2$$

where $y_{i,t}$ denotes the under-1 infant mortality rate in county i and year t, defined as the number of deaths for every 1000 live births before the end of the first year. $x_{i,t}$ includes the percentage of households with piped water from the network, the percentage of households with sewerage connection, the HDI_income variable, the HDI_education variable, and the interaction between HDI_income and the water supply variable. For 1991-2000, we estimate an additional specification that replaces HDI_education with the total fertility rate.[28]

α_i denotes the (unobserved) county effect, which controls for time-invariant sources of unobserved heterogeneity. Without this control, we would expect piped water to be correlated with the error in (1), leading to biased estimates. Indeed, we show this to be the case with a series of cross-sectional regressions below. Examples of unobservables that may be correlated with infant mortality include access to healthcare and breastfeeding behavior.[29] If these variables do not vary in a county over the course of a decade, the county effect will control for them. Similarly, measurement error in infant mortality may vary by county.[30] As long as the rate of measurement error is stable over the course of a decade, α_i will control for its impact on the IMR.

With a constant coefficient vector β and a mean zero restriction on the error term, the typical approach to identifying β with panel data is to estimate the first-differenced model:

[28] Colinearity in these covariates makes it difficult to estimate their distinct effects when they are included within the same regression model.

[29] There are no county-level data maintained on breastfeeding behavior. While there are data available that describe the number of doctors, nurses, and hospitals at the county level, we found that these variables had no explanatory power after controlling for the county effect, α_i.

[30] Victora and Barros (2001), citing Simões (1999), report that the under-registration of infant deaths are estimated to be 52.2% in the North, 66.7% in the Northeast, 6.5% in the Southeast, 13.6% in the South and 23.9% in Center-West, and 43.7% nationally. Most deaths that are not registered occur in the rural areas of the North and Northeast where rates of infectious diseases are higher.

(2) $$y_{i,2} - y_{i,1} = (x_{i,2} - x_{i,1})'\beta + (\varepsilon_{i,2} - \varepsilon_{i,1})$$

by simply regressing the differenced dependent variable on the differenced covariates. Unfortunately, such an approach will not be valid in the quantile regression setting. To see why, we return to the basic model introduced by Koenker and Bassett (1978) and Koenker and Hallock (2001), which allowed marginal effects to vary by quantile. They considered a (cross-sectional) linearly heteroskedastic model of the form:

(3) $$y_i = \alpha_i + x_i'\beta + (x_i'\psi)\varepsilon_i$$

which implies that the θ^{th} conditional quantile of the dependent variable has the following form:

(4) $$\begin{aligned} q_\theta &= \alpha_i + x_i'\beta + x_i'\psi\rho_\theta \\ &= \alpha_i + x_i'(\beta + \psi\rho_\theta) \\ &= \alpha_i + x_i'\beta_\theta \end{aligned}$$

where ρ_θ denotes the θ^{th} quantile of ε_i. We now demonstrate that this model cannot carry through to the panel data model by first-differencing. In the linear heteroskedastic framework, equation (3) becomes:

(5) $$y_{i,2} - y_{i,1} = (x_{i,2} - x_{i,1})'\beta + (x_{i,2}'\psi\varepsilon_{i,2} - x_{i,1}'\psi\varepsilon_{i,1})$$

Taking conditional quantiles of both sides of equation (5) yields:

(6) $$q_\theta(y_{i,2} - y_{i,1} \mid x_{i,1}, x_{i,2}) = (x_{i,2} - x_{i,1})'\beta + q_\theta(x_{i,2}'\psi\varepsilon_{i,2} - x_{i,1}'\psi\varepsilon_{i,1})$$

Since the quantile and difference operators cannot typically be interchanged, the last term in the above expression is not equal to $(x_{i,2} - x_{i,1})'\psi\rho_\theta$. We also note that if we did not allow for the heteroskedastic component, $x_{i,t}\psi$, then the quantile difference function would be a linear function of β plus an additive constant that varied with the quantile. In this restricted setting, marginal effects would not be allowed to vary across quantiles.

We therefore modify the approach described in Chen and Khan (2007). In particular, we impose non-parametric structure on the county effect:

(7) $$\alpha_i = \phi(x_{i,1}, x_{i,2})$$

Where $\phi(\cdot)$ is an unknown function that allows for arbitrary dependence on the covariates. This structure generalizes the typical random effects approach, which does not permit α_i to depend upon covariates. It also generalizes approaches which impose parametric specification on α_i, such as Chamberlain(1982), and Abrevaya and Dahl(2006). Consequently, we have the following functional form for the conditional quantile functions:

(8) $$q_\theta(y_{i,t} \mid x_{i,1}, x_{i,2}) = \phi(x_{i,1}, x_{i,2}) + x_{i,t}'\beta + x_{i,t}'\psi\rho_\theta$$

This implies that the first differences in the conditional quantile functions are of the form:

(9) $$\begin{aligned} q_\theta(y_{i,2} \mid x_{i,1}, x_{i,2}) - q_\theta(y_{i,1} \mid x_{i,1}, x_{i,2}) &= (x_{i,2} - x_{i,1})'\beta + (x_{i,2} - x_{i,1})'\psi\rho_\theta \\ &= (x_{i,2} - x_{i,1})'(\beta + \psi\rho_\theta) \\ &= (x_{i,2} - x_{i,1})'\beta_\theta \end{aligned}$$

which implies an ability to estimate quantile-varying marginal effects. Of course, the above equations do not translate directly into a feasible estimation procedure since the conditional quantile functions are unknown. The approach can be implemented, however, by following a simple two-step procedure:

(i) Non-parametrically estimate the conditional quantile functions $q_\theta(y_{i,2} \mid x_{i,1}, x_{i,2}) \quad t = 1, 2$. Denote these estimated values as $\hat{q}_\theta(y_{i,t} \mid x_{i,1}, x_{i,2})$.

(ii) Regress the differenced fitted values, $\hat{q}_\theta(y_{i,2} \mid x_{i,1}, x_{i,2}) - \hat{q}_\theta(y_{i,1} \mid x_{i,1}, x_{i,2})$ on the differenced regressors, $(x_{i,2} - x_{i,1})$, to estimate β_θ.

As discussed in Chen and Khan (2007), this procedure is very simple to implement, requiring little more than STATA or comparable statistical software.

We implement this panel data procedure for 3 time periods: 1970-1980, 1980-1991, and 1991-2000. For the weighted regressions, we weight the observations by the average county-level population over the two years. We also estimate unweighted regressions in order to check the robustness of our results. Finally, we use 8,000 bootstrap simulations to recover standard errors for our estimates.

5.2 Simulation – Marginal effects of piped water and averted infant deaths

After estimating, we simulate the policymaker's expectation of averted infant deaths resulting from the additional provision of piped water. We make this calculation using the estimates from both the mean and quantile regression specifications. We apply these estimates to a simulated change of one percentage point in the number of households receiving piped water in each county.

Counties are grouped as being high or low in each of these four development indicators: piped water, piped sewage, income and education. The cutoff is at the median of each of these variables.[31] We therefore have 16 groups of counties corresponding to 16 possible combinations of high or low values for the four indicators.[32] We calculate for each group of counties their intra-group

[31] In 1970 and 1980, we use the cutoff of sewage coverage at the 55th and 65th percentile of observations. In those years, a substantial number of counties did not have any households connected to the sewage network.

[32] We limit the number of covariates in part because of the need to create group of counties that are similar in their covariates. With four covariates, we generate 16 groups of counties, with several observations in each group.

mean HDI_income and HDI_education. Next, for each group of counties, we calculate the intra-group distribution of the infant mortality rate. A county will therefore occupy the θ^{th} percentile of the conditional infant mortality rate distribution (i.e., within a group of counties that are similar in their four development indicators).

For each county in a given group, we calculate the marginal effect of piped water on its infant mortality rate (measured as deaths per 1000 live births) using estimates from the appropriate quantile regression and accounting for local conditions as captured by the HDI_income. We then simulate the effect of an increase of one percentage point in the number of households with piped water.[33]

Next, for each county in a given group, we calculate the expected deaths averted from the increase of ten percentage points in the number of households with piped water supply in each county. The expected number of averted deaths in county *i* is given by the marginal effect of water in county *i* on deaths per 1000 live births x the number of expected births in county *i* x (1/1000), where the number of expected births = (total fertility rate in county *i*) x (female population in county *i*). We sum the number of averted deaths across counties to obtain the total number of averted deaths in all Brazil

To estimate the expected deaths averted using the results from the mean regressions, we first calculate the marginal effect of water estimated at the mean. The expected number of averted deaths in county *i* is given by the marginal effect of water in county *i* on (deaths per 1000 live births) x (number of expected births in county *i*) x (1/1000). The total of averted deaths in all Brazil is the sum of the averted deaths across all counties in Brazil.

[33] For counties whose intra-group IMR is below the 10^{th} percentile, we use the estimates from the 10^{th} quantile regression. For counties whose intra-group IMR is between the 10^{th} and 20^{th} percentile, we use the estimates from the 20^{th} quantile regression, and so on. We use the estimates from the 90^{th} quantile, for counties whose intra-group IMR is between the 80^{th} and 90^{th} percentiles, as well as for counties whose intra-group IMR is above the 90^{th} percentile.

6 Results

6.1 Regression results

Table 2 tabulates the results from the regressions weighted by county-level population. Results from the mean regression are in column 1, while those from the quantile regressions are in columns 2 to 10. Panels A, B and C show results from the fixed effect regressions for 1970-1980, 1980-1991, 1991-2000, respectively. In this section, we discuss the signs, sizes, and statistical significance of the coefficients on determinants of IMR other than water across the various specifications.[34] The marginal effect of water, which is dependent on the county-level HDI_income, is discussed in the section 6.2. At the conclusion of section 6.2, we report cross-sectional estimates in order to emphasize the important role being played by our quantile panel data technique.

In all decades, an increase in the percentage of households on the sewage network reduces infant mortality. In 1970-1980 (i.e., when the average county in Brazil was least developed), the estimates are the largest; quantile regressions indicate that a ten percentage point increase in the number of households connected to the sewage network reduces infant mortality by 0.35 to 0.49 deaths per 1000 live births. In contrast, the mean estimates suggest a reduction of only 0.18 deaths per 1000 live births. By 1991-2000, the marginal effect of the sewage network declines relative to its effect two decades earlier. The one percentage point intervention reduces deaths by only 0.04 to 0.17 deaths per 1000 live births. The mean estimates indicate only 0.1 avoided death per 1000 live births.

The marginal impact of education is larger and uniform across the IMR distribution in 1970-1980 and smaller and asymmetric in the latter decades. In 1970-1980, a 0.01 unit increase in education-related HDI reduces infant mortality by 0.22 to 0.29 deaths per 1000 live births. The asymmetry is more pronounced in 1980-1991 than in 1991-2000. In 1980-1991, a 0.01 unit increase in education-related HDI reduces infant mortality by 0.75 deaths in the 10^{th} percentile and by almost two and a half times that amount (i.e. 1.97 deaths) at the 90^{th} percentile.

[34] We discuss statistical significance at and below the 0.10 level.

In 1991-2000, the gap is smaller with 0.78 avoided deaths at the 10th percentile and 1.1 avoided deaths at the 60th percentile and above. In both decades, the mean estimate falls between the estimates from the 10th and 90th quantile regressions.

Table 3 shows the results from the un-weighted regressions. In general, results correspond to those from the weighted regressions.

6.2 Marginal impact of piped water

We report the impact of a one percentage point increase in the number of households with piped water supply. We use the coefficients from the weighted quantile panel data regressions for 1970-1980, 1980-1991 and 1991-2000 (all of which are significant at the 0.05 level), and intra-group means from 1980, 1991, and 2000, respectively. Results using the intra-group mean from 1970, 1980, and 1991 are qualitatively similar.

Figure 1 illustrates the marginal effects of piped water for each group of counties (estimated at the deciles), using results from the weighted quantile regressions. These groups represent counties based on their performance in their development indicators: poor in all four indicators (Grp 1), in three (Grp 2-5), in two (Grp 6-11), in one (Grp 12-15) and none (Grp 16). Table 4 tabulates the development indicators for these groups of counties.

As seen in these sixteen graphs, the impact of water is sizable for many groups of counties in 1970-1980 and for most groups of counties in 1980-1991. In comparing groups of counties within a time period, especially within 1970-1980 and 1980-1991, we see that piped water exerts a greater protective impact on those counties that measure poorly on its development indicator. Consider the worst group of counties (groups 1-3) that perform poorly in all four or in at least three of their development indicators. The largest reductions are 0.45 to 0.56 deaths per 1000 live births for 1970-1980 and 0.75 to 0.86 deaths per 1000 live births for 1980-1991. These reductions are sizable compared with the mean of

125, 87 and 49 deaths per 1000 live births in 1970, 1980 and 1991.[35] In contrast, the reductions in deaths are smaller for the group of counties with the best performance in their development indicators (group 12-16). The largest reductions are only 0.17 to 0.36 deaths per 1000 live births for 1970-1980 and only 0.36 to 0.62 deaths for 1980-1991.

We also find that within a group of counties with similar development indicators, piped water exerts a stronger protective effect at the upper tail of the conditional IMR distribution in 1980-1991 and 1991-2000. The asymmetry across the IMR distribution is largest in 1980-1991. Consider counties in group one. As seen in the first graph, in 1970-1980, additional piped water supply in 1970-1980 reduces the IMR by 0.41 to 0.56 deaths per 1000 across the percentiles. Between 1980 and 1991, the gap across percentiles grows – the additional piped water reduces IMR by 0.37 deaths per 1000 live births at the 10th percentile of the conditional IMR and by 0.86 deaths per 1000 at the 90th percentile. By 1991-2000, the magnitude of the asymmetry declines with the decreasing effectiveness of water. The additional water supply reduces IMR by 0.001 deaths per 1000 live births at the 10th percentile and by 0.2 deaths at the 90th percentile of the conditional IMR. As seen in graphs 2-16, in other groups of counties, the impact of piped water displays the similar pattern of large asymmetry across the IMR distribution in the period 1980-1991 and of smaller gaps across the distribution in the following decade.

In 1970-1980, we actually find that water is less effective for many groups at higher deciles in the conditional IMR distribution. This is surprising given the results in the following two decades, but may be explained by the changing overall level of development in Brazil between 1970 and 2000. In particular, the level of development in the 90th percentile of the conditional IMR distribution in Brazil in 1970 would have been far worse than in the 90th percentile in 2000. That level of development may have been too low for increased water resources

[35] We describe reductions in infant mortality resulting from a one percentage point increase in households with piped water. In practice, the mean increase is about 8.9 percentage points between 1970 and 1980, 18.2 percentage points between 1980 and 1991, and 20.2 percentage points between 1991 and 2000.

to be effective in reducing infant mortality – complimentary public health inputs may have been required to make piped water provision effective. By 1980-1991, those inputs were more likely to be in place, making piped water a particularly effective tool for reducing infant mortality in the 90th decile of the conditional IMR distribution.

The final pattern we see is the declining marginal effects of water by 1991-2000. For the worst group of counties, the largest marginal impact of water is about 0.56 and 0.86 deaths per 1000 live births in 1970-1980 and 1980-1991, respectively. In contrast, the reductions amount to only 0.18 to 0.20 deaths by 1991-2000. While this reduction is not as large as it might seem (given the accompanying reduction in mean IMR over the same period), it is significant. The declining marginal effects of water by 1991-2000 are also evident for the best group of counties. The largest reductions in 1970-1980 and 1980-1991 are about 0.36 deaths and 0.62 deaths per 1000 live births, while the reductions in 1991-2000 are only 0.11 to 0.16 deaths.

Table 5 shows the number of averted deaths as a result of an increase of one percentage point in the number of households with piped water in each county. For the weighted regressions, calculations using the quantile panel data procedure are tabulated in columns (1) and (2), while those using the mean fixed effect regressions are tabulated in columns (3) and (4). Columns (2) and (4) use intra-group means and Brazil-wide means from the year t+1, while column (3) and (4) use the values from year t. The corresponding values for the unweighted regressions are in columns (5)-(8).

Table 5 indicates that in all but one instance, the mean estimates overstate the impact of piped water as measured by the quantile regressions. The divergence between the mean and quantile regressions is largest in the 1980-1991 period. That greater gap is not surprising, given the large asymmetry in the marginal effects of water across the conditional IMR distribution in that period. In 1980-1991, estimates from the weighted quantile regressions suggest 73,000 to 87,000 averted deaths while those from the weighted mean regressions are approximately double those figures. Correspondingly, the divergence between the mean and quantile regression is smallest in the 1991-2000 period. That smaller gap is unsurprising given the small gap in the marginal effects of water across the conditional IMR distribution in that period. Results from the unweighted regressions (columns 5-8) replicate the pattern of the mean estimates, overstating the quantile effects with the largest gap between the two estimates appearing in 1980-1991.[36]

Our simulations indicate that the policymaker would overestimate the number of averted infant deaths should she rely on estimates from the mean regression when new water resources are distributed in proportion with population. Furthermore, additional piped water connections have historically been placed in areas with low conditional IMR, where that provision has less of a protective effect compared to areas with high conditional IMR. This would only further exacerbate the overstatements found in our simulations.

The cross-sectional estimates, as seen in Table 6, suggest that greater provision of piped water is correlated with *larger* infant mortality rates. This counterintuitive result is likely to be an artifact from the systematic underreporting of infant mortality rates in areas on the upper tails of the IMR distribution that tend to receive less water. We see a particularly strong correlation between water supply and increased mortality at the higher conditional quantiles of the IMR, although the size of the bias diminishes in the latter years of the analysis.

7 Discussion and Policy Implication

For those populations at greatest risk, can the provision of piped water reduce the infant mortality rate or are complementary inputs such as income and other public health infrastructure required? Our results are consistent with Shuval et al.'s (1981) threshold-saturation hypothesis, in which the relationship between water supply and IMR varies with changing socioeconomic levels. We find that water has a small effect in the most undeveloped places (i.e., when we

[36] One exception to our conclusions is the larger number of averted deaths from the quantile regressions relative to the mean regressions for 1970. This is likely a result of the reversed asymmetry in the marginal effects across deciles described above for 1970-1980.

look at the high conditional quantiles in 1970-80). As counties start to develop (i.e., the higher quantiles in 1980-91), the protective effect of water on IMR starts to rise rapidly. As counties become more developed (i.e., low quantiles in 1980-91) protective the effect of water declines. Finally, when very developed (i.e., low quantiles in 1991-00), the effect of water is very small.

In 1980-1991, the marginal impact of piped water is greatest in those counties with poorest performance in their observable development indicators. In those counties, in 1980-1991, the largest reduction in deaths from a one percentage point increase in households with piped water is 0.86 deaths per 1000 live births. In contrast, the largest reduction in counties with the best development indicators is only 0.62 deaths per 1000 live births. In addition, among those counties that share common development indicators, we find that piped water exerts a stronger protective effect in those counties that occupy higher positions in the conditional IMR distribution (i.e., counties that are worse in unobservable development indicators), except for 1970-1980 when these counties may have been too undeveloped for piped water to have been effective.

Our results therefore show that (1) piped water provision can cause significant reduction in the IMR (when accompanied by a basic level of other public health inputs); and (2) the impact of a piped water provision policy is determined in large part by how those piped water connections are distributed. Ignoring costs of provision, our results suggest that, from the perspective of health outcomes, new piped water resources should be targeted to the most disadvantaged communities.[37] [38] It is, moreover, not enough to simply consider income, education, and sewage in defining "disadvantage". Policy-makers need

[37]Policymakers may consider other factors in piped water placement such as population density. In the absence of cost data, we cannot provide a cost-benefit analysis on the ideal distribution for additional piped water networks. We acknowledge that the provision of piped water may be cheaper in areas with good development indicators and/or low conditional IMR. These locations may already have a minimal level of existing infrastructure. New outlays of pipelines may have to be undertaken in disadvantaged areas.

[38] "It has been suggested that piped water disproportionately benefit the better-off people of a village" (Mohan, 2005). Further interventions would have to be undertaken to overcome social constraints and connections costs that prevent the vulnerable households from accessing the network.

to look for additional factors (i.e., unobserved factors in our analysis) that make IMR high in allocating water.

Methodologically, these results highlight the importance of applying the quantile regression framework to recover the marginal effects of water at various parts of the conditional distribution of the IMR. Results on the marginal effects at various parts of the distribution differ substantially from that at the mean of distribution. Indeed, focusing on the mean of the distribution can lead to an underestimate of the potential impact of piped water intervention in higher percentiles of the conditional IMR distribution. Our results for piped water intervention correspond with the growing literature on the heterogeneity of program impacts across the quantiles of the conditional distribution of the outcome variable and the insufficiency of mean estimates to represent this policy-relevant heterogeneity.

Quantile estimation for the evaluation of policy is, however, quite difficult. Policies are not often allocated randomly, and good instruments may not be available. Traditional quantile regression is not generally feasible in the panel data context. Our quantile panel data approach, however, can be widely applied to the evaluation of other programs that provide health inputs or public goods in developing countries. This method allows policymakers to understand the impact of these programs on the subpopulations that are most at risk and these subpopulations tend to occupy the tails in the conditional distribution. Amidst the scarcity of random assignment and viable instruments, but with the growing availability of panel data in developing countries, the panel data approach provides a promising strategy to address the issue of bias arising from unobservables, albeit only time invariant ones, within the context of quantile regressions.

Bibliography

Abadie A, Angrist J, Imbens G. Instrumental variables estimates of the effect of subsidized training on the quantiles of trainee earnings. Econometrica 2002; 70(1); 91-117.

Abrevaya J. The effects of demographics and maternal behavior on the distribution of birth outcomes" Empirical Economics 2001; 26 (1); 247-257.

Abrevaya J, Dahl CM. The effects of birth inputs on birthweight: evidence from quantile estimation on panel data. Journal of Business and Economic Statistics 2006 (conditionally accepted).

Alves D, Belluzzo W. Infant mortality and child health in Brazil. Economics and Human Biology 2004; Special Issue 2(3); 391-410.

Anderson JE. Survey of maternal and child health and family planning in Northeastern Brazil: measurement of the duration of breastfeeding and postpartum amenorrhea. US Department of Health and Human Services, Center for Disease Control working paper; 1991.

Arias O, Hallock K, Sosa-Escudero, W. "Individual heterogeneity in the returns to schooling: instrumental variable quantile regression using twins data," Empirical Economics 2001; 26; 7-40.

Aturupane H, Deolalikar AB, Gumewardena D.The determinants of child weight and height in Sri Lanka, A quantile regression approach, paper presented at the WIDER conference on advancing health equality, Helsinki, 29-30 September 2006.

Barnum, H. Interaction of infant mortality and fertility and the effectiveness of health and family planning programs. Policy Planning and Research Working Papers, Population Health and Nutrition, Population and Human Resources Department, the World Bank, WPS 65; July 1988.

Barrera, A. The role of maternal schooling and its interaction with public health programs in child health production. Journal of Development Economics 1990a; 32; 69-91.

Barrera, A. The interactive effects of mother's schooling and unsupplemented breastfeeding on child health. Journal of Development Economics 1990b; 34; 81-98.

Bitler M, Gelbach J, Hoynes H. Distributional impacts of the Self-Sufficiency Project. National Bureau of Economic Research Working Paper 11626; September 2005.

Bitler MP, Gelbach J, Hoynes H. What mean impacts miss: distributional effects of welfare reform experiments. American Economic Review; 2006; 96; 988-1012

Black, RE. Diarrheal diseases and child morbidity and mortality. In: Mosley WH, Chen LC (Eds), Child survival: strategies for research. Population and Development Review 1984; 10 (suppl); 141-161

Black RE, Morris SS, Bryce J. Where and why are 10 million children dying every year? Lancet 2003; 361 (9376); 2226-2234

Bourne PG. Water and sanitation for all. In: Bourne PG (Ed), Water and Sanitation: Economic and Sociological Perspectives, Academic Press; Orlando; p.1-20

Briscoe J. Intervention studies and the definition of dominant transmission routes. American Journal of Epidemiology 1984; 120(3); 449-455

Briscoe J. Evaluating water supply and other health programs: short-run vs long-run mortality effects. Public Health 1985; 99; 142-145

Buchinsky M. Recent advances in quantile regression models: a practical guide for empirical research. Journal of Human Resources 1998; 33(1); 88-126.

Butz WP, Habicht JP, DaVanzo J. Environmental factors in the relationship between breastfeeding and Infant mortality: the role of sanitation and water in Malaysia. American Journal of Epidemiology 1984; 119(4); 516-25,

Cairncross S. Health impacts in developing countries: new evidence and new prospects Journal of the Institution of Water and Environmental Management 1990; 4(6); 571-577

Cairncross S. Handwashing with soap – a new way to prevent ARIs? Tropical Medicine and International Health 2003; 8(8); 677–679.

Chen S, Khan S. Semiparametric estimation of non-stationary censored panel model data models with time-varying factor. Econometric Theory 2007; forthcoming

Chernozhukov V, Hansen C. An IV model of quantile treatment effects. MIT Department of Economics Working Paper 02-07, 2001

Chernozhukov V, Hansen C. An IV model of quantile treatment effects. Econometrica 2005; 73(1); 245-261.
(The published version Econometrica 2005 omits the application example).

Clasen TF, Brown J, Collin, Suntura O , Caincross S. Reducing Diarrhea Through the Use of Household-Based Ceramic Water Filters: A Randomized Controlled Trial in Rural Bolivia. American Journal of Tropical Medicine and Hygiene 2004 70 (6): 651–57.

Conroy RM, Elmore-Meegan M, Joyce T, McGuigan K, Branes J. Solar disinfection of drinking water and diarrhoea in Maasai children: a controlled field trial. Lancet. 1996; 348; 1695–1697.

Crump JA, Otieno P, Slutsker L, Keswick BH, Rosen DH, Hoekstra RM, Vulule JM, Luby SP. Household based treatment of drinking water with flocculant-disinfectant for preventing dirrhoea in areas with turbid source water in rural western Kenya: a cluster randomized trial. British Medical Journal 2005; 331; 478-83.

Curtis V, Kanki B, Mertens T, Traore E, Diallo I, Tall F, Cousens S. Potties, pits and pipes: explaining hygiene behaviour in Burkina Faso. Social Science and Medicine 1995; 41; 383–393.

Cutler D, Miller G. The role of public health improvements in health advances: the twentieth century United States. Demography 2005; 42(1); 1-22

Djebbari H, Smith J. Heterogeneous Program Impacts in PROGRESA, mimeo, Laval University and University of Michigan working paper 2005.

Esrey SA, Feachem RG, Hughes JM. Interventions for the control of diarrhoeal disease among young children: improving water supplies and excreta disposal facilities. Bulletin of the World Health Organization 1985; 63; 757–772

Esrey SA, Habicht JP. Epidemiologic evidence for health benefits from improved water and sanitation in developing countries. Epidemiologic Reviews 1986; 8; 117-127

Esrey SA. Habicht JP. Maternal literacy modifies the effect of toilets and piped water on infant survival in Malaysia. American Journal of Epidemiology 1988; 127(5):1079-1087

Esrey SA, Habicht JP, Casella G. The complementary effect of latrines and increased water usage on the growth of infants in rural Lesotho. American Journal of Epidemiology 1992; 135(6); 659-666

Fitzenberger B, Koenker R, Machado JAF. Economic applications of quantile regression. Studies in empirical economics, Heidelberg; Physica Verlag; 2002

Habicht JP, Da Vanzo J, Butz WP, Mother's milk and sewage: their interactive effects on infant mortality. Pediatrics 1988; 81(3); 456-461

Heckman J, Smith J, Clements N. Making the most out of programme evaluations and social experiments: accounting for heterogeneity in programme impacts. In: The Review of Economic Studies; Special Issue: Evaluation of Training and Other Social Programmes; 1997 64(4); 487-535.

Jalan, J, Ravallion M, Does piped water reduce diarrhea for children in rural India? Journal of Econometrics 2003; 112(1); 153-173

Kremer M, Leino L, Miguel E, Zwane AP Spring cleaning: a randomized evaluation of source water improvement, Department of Economics, Harvard University working paper, 2007

Koenker R, Bassett G. Regression Quantiles. Econometrica 1978; 46; 33(51)

Koenker R, Hallock KF. Quantile regression. Journal of Economic Perspectives 2001; 15 (4); 143-156

Lee LF, Rosenzweig MR, Pitt MM. The effects of improved nutrition, sanitation, and water quality on child health in high-mortality populations. Journal of Econometrics 1997; 77; 209-236

Merrick TW. The effect of piped water on early childhood mortality in urban Brazil: 1970 to 1976. Demography 1985; 22 (1);1-24

Mintz, E, Bartram J, Lochery P, Wegelin M, Not Just a Drop in the Bucket: Expanding Access to Point-of-Use Water Treatment Systems. American Journal of Public Health 2001; 91(10); 1565–1570.

Mobarak AM, Rajkumar AS, Cropper M. The political economy of health-care provision and access in Brazil. World Bank Policy Research Working Paper 3508; 2005

Mohan P. Inequities in coverage of preventive child health interventions: The Rural Drinking Water Supply Program and the Universal Immunization Program in Rajasthan, India. American Journal of Public Health 2005; 95(2); 241-244.

Pebley AR, Millman S. Birthspacing and child survival, International Family Planning Perspectives 1986; 12(3); 71-79

Potter JE, Schmertmann CP, Cavenaghi SM. Fertility and Development: Evidence from Brazil. Demography 2002; 39(4): 739-761

Rosenzweig MR, Schultz TP. Child mortality in Colombia: individual and community effects. Health Policy and Education 1982; 2(3-4):305-348

Rosenzweig, MR, Wolpin, K. Evaluating the Effects of Optimally Distributed Public Programs: Child Health and Family Planning Interventions. American Economic Review 1986; 76 (3); 470-482

Sastry N. Community characteristics, individual and household attributes, and child survival in Brazil. Demography 1996; 33(2); 211-229

Sastry N, Burgard S. Diarrheal disease and its treatment among Brazilian children: stagnation and progress over a ten-year period, RAND Labor and Population Program WPS 02-04; 2002

Senauer B, Kassouf AL. The effects of breastfeeding on health and the demand for medical assistance among children in Brazil. Economic Development and Cultural Change 2000; 48 (4); 719-736

Shuval HR, Tilden RL, Perry BH, Grosse RN. Effect of investments in water supply and sanitation on health status: a threshold-saturation theory. Bulletin of World Health Organization 1981; 59(2); 243-248

Simões CC. Estimativas da mortalidade infantil por microrregiões e municípios. Brasília: Ministério da Saúde (SPS); 1999.

Sinha K. Household characteristics and caloric intake in rural India: a quantile regression approach. Australian National University working paper; 2005

Stefani P, Biderman C. Returns to education and wage differentials in Brazil: a quantile approach. Economics Bulletin 2006; 9(1); 1-6

Thomas D, Strauss J. 1992. Prices, infrastructure, household characteristics and child height. Journal of Development Economics 1992; 39(2); 301-331

Thomas D, Strauss J, Henriques MH. Child survival, height for age and household characteristics in Brazil. Journal of Development Economics 1990; 33(2); 197-234

Thomas D, Strauss J, Henriques MH. How does mother's education affect child height? Journal of Human Resources 1991; 26(2); 183-211

Troesken W. Race Disease and the Provision of Water in American cities (1889-1921), Journal of Economic History 2001; 61(3); 750-756

United Nations. The Millenium Development Goals Report 2005; United Nations: New York; 2005

Victora CG. Potential interventions to improve the health of mothers and children in Brazil. Revista Brasileira de Epidemiologia 2001; 4(1); 3-69 (In Portuguese)

Victora CG, Barros FC. Infant mortality due to perinatal causes in Brazil: trends, regional patterns and possible interventions. Sao Paolo Medical Journal 2001; 191(1); 33-42

Victora CG, Olinto MTA, Barros, FC, Nobre LC. Falling diarrhea mortality in Northeastern Brazil: did ORT play a role? Health Policy and Planning 1996; 11(2); 132-141

Victora CG, Smith PG, Vaughan JP. Social and Environmental Influences on Child Mortality in Brazil: Logistic Regression Analysis from Census Files Journal of Biosocial Science 1986 18(1); 87-101

Victora CG, Smith PG, Vaughan JP, Nobre LC, Lombardi C, Teixeira AMB, Fuchs SC, Moreira LB, Gigante LP, Barros FC. 1987. Evidence for protection by breast-feeding against infant deaths from infectious diseases in Brazil. Lancet 1987; 2(8554); 319–322

Victora CG, Smith PG, Vaughan JP, Nobre LC, Lombard C, Teixeira AMB, Fuchs SC, L.B. Moreira LB, Gigante LP, Barros FC. Water supply, sanitation and housing in relation to the risk of infant mortality from diarrhea. International Journal of Epidemiology 1988; 17(3); 651-654

Watson T. Public health investments and the infant mortality gap: evidence from federal sanitation interventions in US Indian Reservations. Journal of Public Economics 2006; 90 (8-9); 1537-1560

Weinreb AA. First Politics, Then Culture: Accounting for Ethnic Differences in Demographic Behavior in Kenya. Population and Development Review 2001, 27(3); 437-467

Zwane AP, Kremer M. What works in fighting diarrheal diseases in developing countries? A critical review. World Bank Research Observer 2007; forthcoming.

Randomized trials to test the impact of point-of-use water treatment, such as ceramic filters, solar disinfection and chemical disinfectants, find positive health impacts (Clasen et al, 2004; Conroy et al, 1996; Crump et al, 2005).

Table 1: Summary statistics

	Year	Mean	Std. Dev.
Infant mortality rate	1970	125.3	52.7
(in deaths per 1000 live births)	1980	86.8	45.2
	1991	49.2	24.4
	2000	33.7	18.1
Percentage households	1970	15.1	19.5
with piped water (%)	1980	24.0	21.8
(Water)	1991	42.2	23.9
	2000	62.5	20.5
Percentage households	1970	5.3	12.3
with sewage connections (%)	1980	10.6	19.1
(Sewage)	1991	18.0	26.6
	2000	29.4	30.4
Human Development Index	1970	0.22	0.16
Income (income)	1980	0.54	0.27
	1991	0.56	0.10
	2000	0.61	0.10
Human Development Index	1970	0.40	0.14
Education (education)	1980	0.47	0.14
	1991	0.65	0.13
	2000	0.78	0.09
Total fertility rate	1970	5.9	1.5
	1980	4.5	1.4
	1991	3.6	1.2
	2000	2.8	0.7
Female population	1970	12,796	68,110
	1980	16,366	93,898
	1991	20,322	110,204
	2000	23,542	123,214
Population	1970	25,460	132,485
	1980	32,534	182,528
	1991	40,138	211,990
	2000	46,418	235,553

Table 2: The influence of piped water on infant mortality: panel regressions weighted by county-level population

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Method	**Mean**				**Quantile**					
Quantile		10	20	30	40	50	60	70	80	90
Panel A (1970-1980)										
water	-0.45 **	-0.66 **	-0.75 **	-0.77 **	-0.79 **	-0.80 **	-0.82 **	-0.62 **	-0.63 **	-0.72 **
	(0.04)	(0.11)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.11)	(0.12)	(0.16)
sewage	-0.18 **	-0.41 **	-0.41 **	-0.41 **	-0.48 **	-0.49 **	-0.47 **	-0.37 **	-0.35 **	-0.35 **
	(0.03)	(0.13)	(0.12)	(0.13)	(0.14)	(0.15)	(0.16)	(0.16)	(0.20)	(0.27)
income	-52.1 **	-31.0 **	-39.7 **	-44.7 **	-50.6 **	-50.6 **	-57.6 **	-54.9 **	-61.9 **	-83.0 **
	(2.65)	(6.08)	(6.24)	(6.13)	(6.44)	(6.55)	(7.20)	(7.73)	(8.04)	(10.5)
education	-247 **	-223 **	-218 **	-229 **	-226 **	-239 **	-247 **	-285 **	-274 **	-283 **
	(9.80)	(19.4)	(20.7)	(20.1)	(22.0)	(23.5)	(24.7)	(27.1)	(29.4)	(36.8)
water x	0.38 **	0.54 **	0.65 **	0.73 **	0.81 **	0.77 **	0.84 **	0.69 **	0.65 **	1.00 **
income	(0.05)	(0.12)	(0.12)	(0.12)	(0.12)	(0.12)	(0.12)	(0.14)	(0.13)	(0.18)
Panel B (1980-1991)										
water	-1.39 **	-0.94 **	-0.94 **	-1.03 **	-1.17 **	-1.40 **	-1.60 **	-1.73 **	-1.92 **	-1.97 **
	(0.09)	(0.13)	(0.12)	(0.12)	(0.13)	(0.12)	(0.14)	(0.14)	(0.15)	(0.20)
sewage	-0.14 **	-0.12 **	-0.03 **	-0.05 **	-0.09 **	-0.08 **	-0.08 **	-0.05 **	-0.04 **	-0.04 **
	(0.03)	(0.07)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.08)
income	-68.5 **	-38.8 **	-44.1 **	-54.4 **	-70.8 **	-82.0 **	-94.4 **	-108 **	-132 **	-151 **
	(4.47)	(5.07)	(5.61)	(6.11)	(6.11)	(6.14)	(6.79)	(7.14)	(7.18)	(9.47)
education	-137.5 **	-74.6 **	-97.6 **	-107 **	-118 **	-124 **	-130 **	-151 **	-170 **	-197 **
	(5.50)	(5.97)	(6.26)	(6.42)	(6.45)	(6.95)	(7.90)	(8.88)	(9.44)	(12.1)
water x	1.28 **	1.26 **	1.24 **	1.39 **	1.58 **	1.86 **	2.09 **	2.25 **	2.54 **	2.42 **
income	(0.11)	(0.20)	(0.18)	(0.18)	(0.17)	(0.17)	(0.18)	(0.20)	(0.20)	(0.26)

Notes: No obs. 3568.** significant at 0.05 level * significant at 0.10 level

Standard errors from 8000 bootstrap repetitions.

Table 2 (continued): The influence of piped water on infant mortality: panel regressions weighted by county-level population

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Method	**OLS**				**Quantile**					
Quantile		10	20	30	40	50	60	70	80	90
Panel C1 (1991-2000)										
water	-0.12 **	0.06 **	0.06 **	0.01 **	0.00 **	0.05 **	0.02 **	-0.12 **	-0.16 **	-0.43 **
	(0.03)	(0.13)	(0.09)	(0.10)	(0.10)	(0.11)	(0.11)	(0.10)	(0.12)	(0.18)
sewage	-0.06 **	-0.04 **	-0.05 **	-0.05 **	-0.07 **	-0.08 **	-0.08 **	-0.08 **	-0.09 **	-0.17 **
	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)	(0.02)	(0.03)	(0.03)	(0.05)
income	-7.98	-10.5 **	-6.36 **	-15.7 **	-12.2 **	-6.80 **	-9.18 **	-22.9 **	-21.8 **	-40.0 **
	(6.42)	(14.1)	(12.1)	(12.1)	(12.3)	(13.7)	(14.3)	(14.3)	(17.3)	(22.1)
education	-99.7 **	-78.2 **	-82.6 **	-87.0 **	-96.1 **	-104 **	-113 **	-112 **	-111 **	-105 **
	(2.54)	(6.69)	(5.59)	(5.58)	(6.21)	(6.94)	(6.93)	(6.47)	(7.22)	(10.2)
water x	0.02	-0.13 **	-0.15 **	-0.06 **	-0.04 **	-0.12 **	-0.08 **	0.09 **	0.11 **	0.45 **
income	(0.06)	(0.20)	(0.15)	(0.15)	(0.16)	(0.18)	(0.17)	(0.17)	(0.21)	(0.28)
Panel C2 (1991-2000)										
water	-0.25 **	-0.07 **	-0.10 **	-0.07 **	-0.08 **	-0.12 **	-0.20 **	-0.26 **	-0.21 **	-0.24 **
	(0.04)	(0.12)	(0.09)	(0.10)	(0.11)	(0.12)	(0.11)	(0.11)	(0.12)	(0.14)
sewage	-0.10 **	-0.07 **	-0.09 **	-0.10 **	-0.13 **	-0.13 **	-0.13 **	-0.14 **	-0.17 **	-0.13 **
	(0.01)	(0.02)	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)
income	-70.2 **	-53.3 **	-62.5 **	-69.6 **	-72.1 **	-75.7 **	-81.7 **	-92.6 **	-88.8 **	-85.3 **
	(6.50)	(13.8)	(11.6)	(12.3)	(12.1)	(12.7)	(13.3)	(14.6)	(16.4)	(19.1)
total fertility	7.01 **	5.48 **	6.05 **	6.26 **	6.83 **	7.34 **	7.80 **	7.68 **	8.44 **	9.51 **
rate	(0.26)	(0.56)	(0.51)	(0.55)	(0.63)	(0.62)	(0.62)	(0.67)	(0.73)	(0.88)
water x	0.01	-0.08 **	-0.05 **	-0.13 **	-0.12 **	-0.08 **	0.01 **	0.06 **	-0.02 **	-0.06 **
income	(0.06)	(0.21)	(0.16)	(0.17)	(0.18)	(0.19)	(0.19)	(0.19)	(0.20)	(0.24)

Notes: No obs. 3568. ** significant at 0.05 level * significant at 0.10 level
Standard errors from 8000 bootstrap repetitions.

Table 3: The influence of piped water on infant mortality rates: unweighted panel regressions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Method	**Mean**				**Quantile**					
Quantile		10	20	30	40	50	60	70	80	90
Panel A (1970-1980)										
water	-0.92 **	-0.85 **	-0.99 **	-0.98 **	-1.05 **	-1.07 **	-1.06 **	-0.95 **	-0.88 **	-1.00 **
	(0.05)	(0.09)	(0.08)	(0.09)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.13)
sewage	-0.48 **	-0.36 **	-0.48 **	-0.45 **	-0.50 **	-0.48 **	-0.46 **	-0.44 **	-0.43 **	-0.53 **
	(0.06)	(0.07)	(0.06)	(0.06)	(0.07)	(0.07)	(0.07)	(0.07)	(0.10)	(0.12)
income	-74.7 **	-52.1 **	-63.4 **	-65.9 **	-75.8 **	-78.4 **	-80.5 **	-79.2 **	-83.2 **	-93.8 **
	(2.69)	(4.09)	(3.94)	(3.99)	(4.06)	(4.18)	(4.56)	(4.84)	(4.78)	(6.50)
education	-156 **	-132 **	-127 **	-150 **	-129 **	-137 **	-155 **	-189 **	-195 **	-208 **
	(8.93)	(15.3)	(16.4)	(14.6)	(15.0)	(16.1)	(17.7)	(18.7)	(19.1)	(25.5)
water x	1.12 **	0.88 **	1.12 **	1.15 **	1.28 **	1.26 **	1.26 **	1.22 **	1.12 **	1.32 **
income	(0.06)	(0.10)	(0.09)	(0.09)	(0.08)	(0.08)	(0.09)	(0.09)	(0.10)	(0.14)
Panel B (1980-1991)										
water	-1.35 **	-0.67 **	-0.72 **	-0.89 **	-1.11 **	-1.29 **	-1.53 **	-1.56 **	-1.70 **	-1.94 **
	(0.08)	(0.08)	(0.09)	(0.12)	(0.11)	(0.11)	(0.13)	(0.12)	(0.14)	(0.20)
sewage	0.02 **	-0.08 **	-0.01 **	-0.03 **	-0.02 **	0.01 **	0.03 **	0.08 **	0.16 **	0.19 **
	(0.04)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)	(0.04)	(0.06)
income	-89.0 **	-31.5 **	-38.1 **	-49.9 **	-66.4 **	-78.6 **	-93.6 **	-106 **	-125 **	-149 **
	(3.64)	(3.93)	(4.52)	(5.48)	(5.42)	(5.22)	(6.06)	(5.75)	(6.25)	(8.47)
education	-131 **	-75.6 **	-98.7 **	-106 **	-115 **	-120 **	-128 **	-156 **	-178 **	-204 **
	(4.87)	(5.15)	(5.50)	(6.34)	(6.39)	(6.54)	(7.58)	(9.28)	(9.09)	(12.8)
water x	1.76 **	0.76 **	0.89 **	1.15 **	1.42 **	1.65 **	1.99 **	2.13 **	2.33 **	2.53 **
income	(0.11)	(0.11)	(0.12)	(0.14)	(0.13)	(0.14)	(0.16)	(0.15)	(0.17)	(0.23)

Notes: No obs. 3568. ** significant at the 5% level. * significant at the 10% level
Standard errors from 8000 bootstrap repetitions.

Table 3 (continued): Results for unweighted regressions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Method	**OLS**				**Quantile**					
Quantile		10	20	30	40	50	60	70	80	90
Panel C1 (1991-2000)										
water	-0.28 **	-0.19 **	-0.15 **	-0.18 **	-0.24 **	-0.25 **	-0.33 **	-0.39 **	-0.32 **	-0.38
	(0.04)	(0.08)	(0.06)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)	(0.08)	(0.11)
sewage	-0.05 **	-0.01 **	-0.04 **	-0.03 **	-0.05 **	-0.04 **	-0.05 **	-0.03 **	-0.03 **	-0.04
	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)	(0.04)
income	-28.6 **	-25.5 **	-18.3 **	-25.8 **	-25.3 **	-23.1 **	-26.4 **	-38.5 **	-36.6 **	-51.2
	(5.96)	(12.6)	(10.2)	(10.8)	(11.2)	(11.7)	(12.6)	(11.9)	(14.2)	(19.9)
education	-90.3 **	-71.4 **	-77.4 **	-79.8 **	-88.7 **	-93.9 **	-99.7 **	-102 **	-103 **	-107
	(2.44)	(5.16)	(4.64)	(4.94)	(5.10)	(5.83)	(6.17)	(5.19)	(5.79)	(8.45)
water x	0.32 **	0.25 **	0.17 **	0.20 **	0.30 **	0.31 **	0.41 **	0.48 **	0.34 **	0.40
income	(0.06)	(0.13)	(0.10)	(0.11)	(0.11)	(0.11)	(0.12)	(0.12)	(0.14)	(0.19)
Panel C2 (1991-2000)										
water	-0.37 **	-0.19 **	-0.28 **	-0.31 **	-0.31 **	-0.39 **	-0.47 **	-0.50 **	-0.41 **	-0.45
	(0.04)	(0.08)	(0.07)	(0.07)	(0.08)	(0.08)	(0.07)	(0.08)	(0.09)	(0.11)
sewage	-0.10 **	-0.05 **	-0.08 **	-0.09 **	-0.11 **	-0.12 **	-0.10 **	-0.10 **	-0.12 **	-0.04
	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)
income	-72.8 **	-52.5 **	-56.9 **	-65.6 **	-70.5 **	-72.5 **	-83.3 **	-91.7 **	-94.6 **	-94.3
	(6.13)	(12.2)	(10.2)	(10.7)	(10.9)	(11.2)	(11.6)	(12.4)	(14.9)	(17.9)
total fertility	6.65 **	5.34 **	5.53 **	5.70 **	6.25 **	6.87 **	7.16 **	6.95 **	7.85 **	9.02
rate	(0.26)	(0.52)	(0.46)	(0.50)	(0.59)	(0.55)	(0.50)	(0.57)	(0.65)	(0.82)
water x	0.27 **	0.10 **	0.20 **	0.22 **	0.22 **	0.33 **	0.42 **	0.42 **	0.32 **	0.29
income	(0.07)	(0.14)	(0.11)	(0.12)	(0.12)	(0.13)	(0.12)	(0.13)	(0.16)	(0.19)

Notes: No obs. 3568. ** significant at the 5% level. * significant at the 10% level
Standard errors from 8000 bootstrap repetitions.

Table 4: Counties' groups' performance in development indicators

Group	Water	Sewage	Income	Education
1	0	0	0	0
2	1	0	0	0
3	0	1	0	0
4	0	0	1	0
5	0	0	0	1
6	1	1	0	0
7	1	0	1	0
8	1	0	0	1
9	0	1	1	0
10	0	1	0	1
11	0	0	1	1
12	0	1	1	1
13	1	0	1	1
14	1	1	0	1
15	1	1	1	0
16	1	1	1	1

Note: 1 indicates county scores above the median, and 0 otherwise

Table 5: Estimated no. of averted deaths from a one percentage point increase in households with piped water

Source of coefficients	Regression models	(1) Weighted	(2)	(3)	(4)	(5) Unweighted	(6)	(7)	(8)
		Quantile	Quantile	Mean	Mean	Quantile	Quantile	Mean	Mean
	Year for HDI inc	t1	t2	t1	t2	t1	t2	t1	t2
Panel A	t1=1980 t2=1970	45,000	130,000	64,000	99,000	46,000	170,000	81,000	180,000
Panel B	t1=1991 t2=1980	87,000	73,000	150,000	180,000	76,000	60,000	83,000	100,000
Panel C1	t1=2000 t2=1991	15,000	18,000	23,000	25,000	17,000	23,000	18,000	23,000
Panel C2	t1=2000 t2=1991	44,000	47,000	52,000	56,000	44,000	51,000	43,000	50,000

Notes: Coefficients for the weighted and unweighted regressions are from table 2 and table 3, respectively

Table 6: Correlation of piped water supply and infant mortality rates: cross-sectional regressions (weighted by county population)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Method	OLS				Quantile					
Quantile		10	20	30	40	50	60	70	80	90
Panel A: 1970										
water	0.57 **	0.32 **	0.26 **	0.26 **	0.22 **	0.24 **	0.30 **	0.43 **	0.73 **	2.45 **
	(0.07)	(0.12)	(0.15)	(0.16)	(0.17)	(0.19)	(0.25)	(0.31)	(0.49)	(0.76)
sewage	-0.26 **	-0.46 **	-0.33 **	-0.20 **	-0.03 **	0.01 **	0.11 **	0.17 **	0.20 **	-0.08 **
	(0.05)	(0.11)	(0.11)	(0.13)	(0.16)	(0.21)	(0.30)	(0.42)	(0.58)	(0.71)
income	37.20 **	12.74 **	30.66 **	39.57 **	35.78 **	34.97 **	48.76 **	53.47 **	83.31 **	87.32 **
	(7.31)	(14.44)	(20.01)	(20.40)	(20.93)	(24.96)	(32.32)	(39.35)	(43.41)	(57.41)
education	-264.0 **	-184.8 **	-209.0 **	-234.6 **	-253.5 **	-267.7 **	-277.9 **	-300.8 **	-343.4 **	-401.9 **
	(8.18)	(10.99)	(14.72)	(16.25)	(18.23)	(16.74)	(22.41)	(27.39)	(34.11)	(60.11)
water x	-0.39 **	0.42 **	0.18 **	0.05 **	0.09 **	0.00 **	-0.35 **	-0.66 **	-1.31 **	-2.88 **
income	(0.11)	(0.25)	(0.32)	(0.33)	(0.32)	(0.34)	(0.40)	(0.50)	(0.73)	(1.07)
Panel B: 1980										
water	0.77 **	0.15 **	-0.07 **	0.07 **	0.22 **	0.37 **	0.53 **	0.94 **	1.21 **	1.49 **
	(0.10)	(0.19)	(0.19)	(0.20)	(0.25)	(0.28)	(0.38)	(0.39)	(0.37)	(0.60)
sewage	-0.28 **	0.01 **	-0.03 **	-0.19 **	-0.25 **	-0.32 **	-0.43 **	-0.36 **	-0.33 **	-0.23 **
	(0.03)	(0.07)	(0.10)	(0.13)	(0.13)	(0.14)	(0.14)	(0.14)	(0.12)	(0.25)
income	-34.00 **	-26.60 **	-46.08 **	-44.56 **	-41.23 **	-40.22 **	-41.30 **	-34.56 **	-19.43 **	30.14 **
	(4.82)	(12.58)	(13.60)	(11.42)	(10.62)	(11.24)	(11.79)	(16.67)	(25.81)	(71.48)
education	-187.0 **	-94.50 **	-100.1 **	-129.9 **	-150.5 **	-177.7 **	-182.5 **	-194.8 **	-241.9 **	-299.8 **
	(8.26)	(14.72)	(16.65)	(15.34)	(15.98)	(16.27)	(19.73)	(28.16)	(33.21)	(55.10)
water x	-0.02	0.20 **	0.60 **	0.65 **	0.55 **	0.47 **	0.37 **	-0.18 **	-0.57 **	-1.35 **
income	(0.11)	(0.25)	(0.30)	(0.32)	(0.35)	(0.38)	(0.45)	(0.47)	(0.39)	(0.67)

Notes: No obs. 3658. Constant incl in regressions. (**) significant at the 0.05 level. (*) significant at the 0.10 level.
Standard errors from 8000 bootstrap repetitions.

Table 6 (continued): Correlation of piped water supply and infant mortality rates: cross-sectional regressions (weighted by county population)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Method	OLS				Quantile					
Quantile		10	20	30	40	50	60	70	80	90
Panel C: 1990										
water	-0.66**	-0.83 **	-0.71 **	-0.88 **	-0.96 **	-0.88 **	-0.82 **	-0.68 **	-0.60 **	-0.74 **
	(0.05)	(0.16)	(0.19)	(0.22)	(0.21)	(0.19)	(0.16)	(0.15)	(0.15)	(0.19)
sewage	-0.11**	-0.04 **	-0.03 **	-0.09 **	-0.09 **	-0.11 **	-0.15 **	-0.20 **	-0.23 **	-0.26 **
	(0.01)	(0.02)	(0.03)	(0.04)	(0.03)	(0.05)	(0.06)	(0.06)	(0.05)	(0.07)
income	-125**	-122.0 **	-95.50 **	-112.0 **	-121.0 **	-123.0 **	-141.0 **	-112.0 **	-126.0 **	-167.0 **
	(7.33)	(15.72)	(16.88)	(19.40)	(19.72)	(25.32)	(26.26)	(28.31)	(31.16)	(31.35)
education	-91.5**	-63.63 **	-82.44 **	-93.13 **	-90.13 **	-97.58 **	-93.05 **	-100.2 **	-100.3 **	-80.61 **
	(3.74)	(7.50)	(10.21)	(11.61)	(12.41)	(15.07)	(14.59)	(13.61)	(13.76)	(16.97)
water x	1.33**	1.48 **	1.24 **	1.59 **	1.72 **	1.63 **	1.63 **	1.34 **	1.29 **	1.49 **
income	(0.08)	(0.23)	(0.29)	(0.34)	(0.34)	(0.31)	(0.27)	(0.26)	(0.30)	(0.35)
Panel D: 2000										
water	-0.56**	-0.66 **	-0.76 **	-0.81 **	-0.84 **	-0.81 **	-0.75 **	-0.57 **	-0.34 **	-0.40 **
	(0.04)	(0.18)	(0.18)	(0.23)	(0.22)	(0.18)	(0.14)	(0.17)	(0.17)	(0.21)
sewage	-0.07**	-0.06 **	-0.07 **	-0.08 **	-0.06 **	-0.08 **	-0.10 **	-0.10 **	-0.13 **	-0.03 **
	(0.01)	(0.02)	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)	(0.05)	(0.05)	(0.07)
income	-124**	-115.0 **	-127.3 **	-131.8 **	-130.4 **	-128.5 **	-136.9 **	-126.0 **	-104.0 **	-144.8 **
	(6.65)	(13.41)	(13.88)	(17.33)	(20.59)	(23.48)	(24.42)	(27.45)	(26.60)	(32.98)
education	-96.0**	-55.11 **	-69.77 **	-80.88 **	-95.01 **	-100.5 **	-89.98 **	-86.41 **	-111.1 **	-100.6 **
	(4.10)	(13.07)	(14.10)	(14.53)	(14.95)	(14.15)	(18.40)	(20.41)	(20.97)	(33.39)
water x	1.14**	1.14 **	1.31 **	1.42 **	1.49 **	1.46 **	1.40 **	1.12 **	0.86 **	0.96 **
income	(0.07)	(0.29)	(0.30)	(0.38)	(0.35)	(0.29)	(0.23)	(0.25)	(0.26)	(0.33)

Notes: No obs. 3658. Constant incl in regressions. (**) significant at the 0.05 level. (*) significant at the 0.10 level.
Standard errors from 8000 bootstrap repetitions.

Figure 1 – Effects of 1 Percentage Point Increase in Piped Water by Year, Development Group, and Conditional IMR Decile (Quantile Panel Data Estimates)