Human Capital and Infectious Disease:
Evidence from Mexico's Clean Water Program

4 March 2019

Sonia R. Bhalotra
University of Essex, UK
srbhal@essex.ac.uk

Ryan Brown
University of Colorado Denver, USA
ryan.p.brown@ucedenver.edu

Atheendar S. Venkataramani
University of Pennsylvania, USA
atheenv@pennmedicine.upenn.edu

Abstract: More than three-quarters of a billion people live without a close source of clean water. While, the immediate impact of clean water access on infant mortality is well documented, there is very limited evidence on the long-term effect of chlorinated water. We exploit exogenous variation created by the implementation of a major clean water reform in Mexico in 1991, Programa de Agua Limpia (PAL), to investigate the impact of exposure to chlorinated water early in life on cognitive and physical development. We estimate that experiencing a one standard deviation reduction in childhood diarrhea mortality rates from PAL throughout infancy leads to ~6% increase in cognitive assessment score and .15 standard deviation increase in height in adolescence. In early adulthood, the effects on human capital persist, and lead to increased hourly earnings.

Keywords: clean water, diarrhea, cognitive development, height, Mexico
JEL codes: I15, I38, I12, I14, H51

Acknowledgements: This paper has benefited from presentation at the AEA meetings in San Diego, the Colombian Economics Conference, and at seminars at the Paris School of Economics, the Institute of Fiscal Studies in London, Montana State University, and the Universities of Pompeau Fabra, Bristol, Oxford, Sussex, Essex. We received helpful comments from Orazio Attanasio, Manuel Bagues, Erlend Berg, Antonio Ciccone, Patricia Cortes, James Fenske, Paul Schultz, Marcos vera Hernandez, Daniel Rees, and Simon Quinn. Damian Clark, Peter Gerrish, Claudia Herresthal, Salomo Hirvonen, and Yichau Wu provided excellent research assistance. All errors are our own.

I. Introduction

The United Nations and WHO have declared access to clean water a basic human right. Having water to drink is fundamental to human existence and exposure to unclean water leads to dire health consequences such as the spread of diarrhea, cholera, dysentery, typhoid and polio (WHO 2016). For example, it is estimated that each year there are more than 2 million deaths from waterborne diseases. Of these maladies, diarrhea is the most prevalent, and represents the second leading cause of child death in the world (Fischer Walker et al 2013).

This environmental reality is, on the one hand, a rather homogenous phenomenon in that it is an issue that is present throughout almost the entire developing world (UNICEF and WHO, 2017). On the other hand though, access to clean water is quite heterogeneously related to poverty as very few people that lack clean water live in a developed country. This critical difference in environmental risk factors is an important contributing cause of the significant gap that exists in the prevalence of infectious diseases in developing versus developed countries.[1]

The immediate impact that between-country variation in exposure to tainted water can have is well established. In particular there is a large literature that finds, across various contexts and time periods, increased access to clean water significantly lowers infant and child mortality (Argentina: Galiani et al., 2005; Bangladesh: Field et al., 2011; Brazil: Gamper-Rabindran et al., 2005; Mexico: Bhalotra et al., 2017 USA: Cutler and Miller, 2005 and Watson, 2005; among others). While this set of research has helped prompt large investments by governments and NGOs in safe water provision, an estimated 844 million people still do not have a readily available source of clean and protected water (UNICEF and WHO, 2017). Moreover, there is good reason to believe that by focusing only on the immediate mortality effect, researchers and policymakers may be severely underestimating the true value of clean water.

---

[1] The number of healthy life years lost to infectious diseases was 15 fold higher in developing versus developed countries in 2014 (WHO, 2014).

There is a growing consensus that health shocks in the first 1000 days of life can have a persistent impact on the individual's physical and cognitive development (Currie and Vogl, 2013; Almond and Currie, 2011; Cuhna and Heckman, 2007; Heckman 2007; among others). In addition, a robust literature has demonstrated the positive association between health capital, such as stature and intelligence, and economic outcomes (Lin et al., 2018; Glewwe et al., 2017; LaFave and Thomas, 2017; Vogl, 2014; Huang et al., 2013; Case and Paxson, 2008; Heckman et al., 2006; Strauss and Thomas, 1997; among others). These two widely accepted facts suggest that increased access to clean water at birth, which will reduce the prevalence of severe or repeated infections that can divert nutrients away from physical and neurological development during the critical early life period, may have a long-lasting impact on the economy.[2]

Recognition of these mechanisms in the medical literature has generated the provocative suggestion that the much higher prevalence of infectious disease in poor countries may help explain their weaker performance on international intelligence tests and stagnated rate of economic growth. These propositions potentially have major implications for understanding human growth and economic development, standing to inform, for instance, debates in the economics literature concerning the role of innovation-led declines in infectious disease in explaining living standards (Acemoglu and Johnson 2007; Sachs and Malaney 2002; Bleakley 2010; Costa 2013). However, the evidence available to support a causal link between early life exposure to clean drinking water and markers for long-term human capital accumulation is at best suggestive.

The goal of this paper is to provide causal evidence of the contemporary relationship between improved access to clean water in infancy and physical and cognitive growth. We focus on infant exposure because the caloric requirements for brain development are higher during this period than at any other stage of the life course,

---

[2] During infancy about 85% of calorie intake is used for brain development (Eppig et al 2010). In addition, the release of inflammatory molecules during an infection may directly impact the developing brain by changing the expression of genes involved in the development of neurons and the connections between them (Deverman and Patterson, 2012).

making infants especially vulnerable to diminished cognitive endowments as a result of diarrhea, a highly morbid and often recurrent symptom of waterborne infections that can severely compromise nutrition, physical growth, and mental development (Fischer Walker et al 2012a; Fischer Walker 2013).

To account for selectivity into access to clean water, we exploit the introduction of the *Programa de Agua Limpia* (National Clean Water Program; PAL) in Mexico, a large-scale nationwide effort. This program provides quasi-random temporal variation, as its introduction was a sudden reaction to the unanticipated threat of cholera created by an epidemic spreading through the countries neighboring Mexico and led to rapid and sizeable drops in childhood diarrhea mortality rates (Bhalotra et al, 2017). PAL also provides rich geographic variation as its impact was systematically larger in areas with worse pre-program water quality, and thus, areas with higher pre-program childhood diarrhea mortality rates (Bhalotra et al., 2017). Our identification strategy exploits both municipality-level pre-program childhood diarrhea mortality rate and birth cohort variation, combining the sharp convergence across Mexican municipalities with the timing of the nationwide reform.

We find sizeable positive effects of exposure to clean water in infancy on performance in Raven's cognition tests and height measured at the ages of 10-16. We estimate that experiencing the average diarrheal mortality decline over our sample period throughout the entire infancy period leads to a ~6% increase in cognitive assessment scores and a .15 standard deviation increase in height during adolescence and that similarly sized effects persist into early adulthood. We also find that the gains to stature are shared equally between boys and girls, but the improvement in cognition, while still present for boys, are stronger for women. In addition we report that, in adolescence, boys with more intense PAL exposure throughout infancy achieved greater attained education, partly through a reduction in grade repetition and by delaying entry into the labor market. Lastly, when exploring economic outcomes in early adulthood, 17-26 years old, we find evidence that early life access to clean water resulted in significant gains to productivity, as measured by earnings per hour.

To provide faith in the causal nature of our estimates we show that the results are not driven by unobserved trends through the inclusion of municipality specific time-varying controls, municipality specific linear time trends, alternative disease mortality rates that capture the general health environment, as well as, by providing an event study analysis. Lastly, to avoid any potential bias from selective mortality, selective fertility and/or selection into live birth we conduct the analyses only using within-family variation.

Our study makes substantive contributions in several domains. We provide the first causal evidence that the benefits of clean water interventions that limit early life exposure to waterborne diseases in a developing country context extend beyond lowering morbidity and mortality to enhancing cognitive skill formation and health[3]. Second, while most research on the causal impacts of early life infectious disease on human capital are restricted to a specific sub-population at one moment in time, we are able to study a nationwide change in the health environment using nationally representative data and investigate whether the benefits of this intervention persists over time. Third, we can explore whether the cognitive and physical gains from early-life exposure to a clean water intervention leads to changes in economic outcomes in early adulthood. Lastly, by being able to directly test for changes in parental composition resulting from the transformation of the health environment, we show that a failure to control for selection bias can substantially alter the magnitude of the estimates.

The remainder of the paper is laid out as follows. Section II discusses the National Clean Water Program, establishes short-run (or "first-stage") impacts on diarrheal disease risk, and describes the limited current evidence regarding early-life access to clean water

---

[3] An association between the number of diarrheal episodes in early childhood and poorer cognitive development has been demonstrated in the medical literature (Fischer Walker et al 2012a; Guerrant, et al, 1999; Niehaus et al 2002) but these studies do not account for the endogeneity of diarrheal episodes beyond covariate adjustment.

and human capital accumulation. Section III discusses the empirical strategy and Section IV the data. The results are discussed in Section V. Section VI concludes.

II. Background

IIa. History and Description of Mexico's National Clean Water Program

Through the late 1980s, infectious diseases were responsible for a significant proportion of infant and child deaths in Mexico. Diarrheal disease was a particularly important scourge, accounting for nearly a quarter of under-5 deaths in this period (Gutierrez et al 1996).[4] Since 1978-1997, public sector efforts have been highly effective in reducing infectious disease deaths (Frenk et al 2003), with credit typically assigned to expansions in access to clean water, sanitation, vaccines and oral rehydration therapy (Sepulveda et al 2007).

This paper focuses on the early 1990s, when the intensity of public health activity increased considerably due to the emergence of a cholera epidemic in other parts of Central and South America. Fears of the outbreak extending to Mexico prompted public health officials to proactively undertake improvements in access to potable water and sanitation and information campaigns to educate local leaders and constituents about cholera and encourage preventative behavior (Gutierrez et al 1996; Sepulveda et al 2006; Sepulveda et al 2007). In April 1991, Mexico implemented a National Clean Water Program (*Programa de Agua Limpia* or PAL). PAL did not expand existing piped water infrastructure or sewage networks, but rather increased water disinfection through existing water pipe infrastructure. To improve water quality in areas without piped water coverage, chloride tablets were disseminated to households and monitoring for commercial bottled water and ice was expanded. Lastly, there was a concerted effort to reduce the use of wastewater in irrigation. The total outlay for the program over the period 1991-1994 was 1 billion USD.[5] Implementation was rapid: within the first eight months of the program the

---

[4] See *Figure 1* for a map of diarrhea prevalence across the Mexican municipalities.

[5] Personal communication with Dr. Jaime Sepulveda.

share of the population receiving chlorinated water rose from 54% (in April 1991) to 85% (at the end of the year), see *Figure 2a*. Similarly, farmland area irrigated with wastewater declined markedly between April and December 1991, see *Figure 2b*.[6] Despite these changes, there is no evidence that PAL crowded out or in other household or public investment in water quality or health infrastructure (Bhalotra et al., 2017).

IIb. Establishing Program Impacts on Diarrheal Disease Mortality

Previous explorations of PAL's impact have hinted that the program had a substantial impact on the citizen's health. Specifically, Gutierrez et al. (1996), Velazquez et al. (2004), and Sepulveda et al. (2006) provide descriptive evidence of large declines in childhood diarrheal disease mortality rates beginning in 1991.[7] These studies, though, only analyze national trends over time, which may reflect other unobserved factors leading to an improved health environment unrelated to PAL. In this section we formalize and restate the evidence from Bhalotra et al. (2017) which uses *Mexican Vital Statistics* data to show there was a trend break and convergence in diarrheal disease mortality after 1991 using age, gender and disease-specific mortality rates by municipality and year.

To estimate the impact of PAL on diarrheal mortality rates among children under-5, Bhalotra et al. (2017) test for differential trend breaks in diarrheal disease mortality rates in 1991 (when PAL was implemented) relative to a control disease (respiratory infections) which should not be directly affected by PAL.[8] They choose respiratory infections because they were the second-leading cause of child mortality in Mexico prior to PAL and because they share several common risk factors with diarrheal diseases (respiratory infections are

---

[6] Attention to water reform is pertinent as not all countries have seen secular progress in provision. Between 1990 & 2005, the percentage of urban households with piped water declined from 50 to 39 in 32 West African countries (World Bank).

[7] Gutierrez et al. (1996) investigates aggregate morbidity rates and finds that from 1990 to 1993 the average number of annual episodes of diarrheal disease morbidity among children decreased from 4.6 to 2.2. Velazquez et al. (2004) report that between 1990-1995 diarrheal disease morbidity in Mexico declined by over 63%.

[8] An equivalent strategy was used in Jayachandran, Lleras-Muney and Smith (2010).

spread through oral droplets and diarrheal diseases are spread by fecal-oral contamination). Moreover, respiratory diseases are an apt comparison as they respond to water quantity, as opposed to diarrheal diseases, which respond to water quality (Fischer Walker et al. 2013b).

Figure 3 provides the descriptive evidence of the impact of PAL. It showing that child mortality from diarrhea was relatively stable between 1985 and 1990 but started to drop dramatically in 1991, leveling off after 1992, consistent with the timing of the National Clean Water Program, while respiratory and vaccine preventable disease mortality rates show a more gradual decline.

To test this formally Bhalotra et al. (2017) estimate the following event study specification over the period 1985-1995:

$$(1) \quad M_{djt} = \alpha_d + \sum_{t=1986}^{1995} \alpha_t (1(Diarrhea_d) * 1(Year = t)) + \sum_{t=1986}^{1995} \mu_t (1(Year = t)) + \lambda_t + \varepsilon_{djt}$$

In equation (1), M_{djt} is the inverse hyperbolic sine transformation of the mortality rate for disease, *d*, in municipality, *j*, and year, *t*.[9] $1(Diarrhea_d)$ is an indicator variable that equals 1 if the cause of death is a diarrheal disease versus the control disease and 1(Year=t) is an indicator variable which equals 1 for observations in year t. Lastly, municipality fixed effects are denoted as λj. Estimates of α_t provide the average differential percentage change in the diarrheal disease mortality rate vs. the control disease mortality rate in year t relative to the baseline year (in this specification, 1985).[10]

The estimates of α_t obtained by calculating equation (1) are found in *Figure 4*. Confirming the descriptive evidence in *Figure 3*, *Figure 4* shows that, as compared to under-5 mortality from respiratory disease, there was a discrete increase in the decline of

---

[9] Estimates obtained using the inverse hyperbolic sine transformation can be interpreted in the same manner as those obtained using a natural logarithm transformation of the dependent variable, with the advantage of being defined at zero (Burbridge, Magee and Robb 1988).

[10] Standard errors are clustered at the municipality level and municipality weights based on the average number of annual pre-intervention live births are used.

under-5 diarrheal mortality with the implementation of PAL in 1991, which persisted throughout the post-implementation years.[11]

Subsequently to estimate the average program effects of PAL during the study period Bhalotra et al. (2017) use a specification of the following form:

$$\begin{aligned} M_{djt} &= \beta_0 + \beta_1(1(Diarrhea_d)*1(Post_t)*Year_t) + \beta_2(1(Diarrhea_d)*1(Post_t)) \\ (2) \quad &+\beta_3(1(Diarrhea_d)*Year_t) + \beta_4(1(Diarrhea_d) + \sum_{t=1986}^{1995} \mu_t(1(Year=t)) + \lambda_t + \varepsilon_{djt} \end{aligned}$$

In equation (2), *1(Post)* represents an indicator variable for post-PAL years (1991 and later), and all other variables are defined as in equation (1). β_1 and β_2, measure the trend and level breaks in 1991 relative to the control disease mortality rates. *1(Diarrhea$_d$)*1Year$_t$* controls for pre-existing trends in diarrheal disease mortality relative to respiratory disease, the year fixed effects control for municipality-invariant differences in the health environment across time, and *1(Diarrhea$_d$)** captures time-invariant differences between diarrheal and control disease mortality rates.

The results of estimating equation (2) on mortality rates among children under-5 (altogether, as well as separately for neonates (0-1 month), postneonates (1-12 months), and children ages 1-4), can be found in Table 1. In all cases there is evidence of statistically significant level and trend break in diarrheal disease mortality as a result of the introduction of PAL. Specifically, with regard to under-5 mortality rates (first column) the results imply that PAL led to a 48% decline in diarrheal disease mortality rates by 1995. Moreover, this decline was not uniform across the country. As shown in *Figure 5* the decline in diarrheal disease mortality occasioned by the Clean Water Program was larger in areas with higher pre-intervention rates.

---

[11] Morbidity from diarrhea declined in line with mortality. The number of diarrheal episodes per infant per year nationwide is estimated to have declined by over 50% from the late 1980s to mid 1990s, with the bulk of that decrease occurring during 1991-1993 (Gutierrez et al 1996).[11] However, there are no region-specific yearly data for morbidity for any duration of time so we follow the convention of using mortality rates as a proxy for morbidity (Bozzoli et al 2009).

IIc. The Long-term Impact of Early Life Access to Clean Water: Previous Evidence in the Economics Literature

There is established inverse relationship between access to clean water and infant and child mortality (Bhalotra et al., 2017; Cutler and Miller, 2005; Field et al., 2011; Galiani et al., 2005; Gamper-Rabindran et al., 2005; Watson, 2005; among others). With few exceptions, though, the research on access to clean water in early life has focused only on the immediate health-related benefits. Specifically, Zhang and Xu (2016) and Beach et al. (2016) represent the totality of the economics literature that claims to investigate the persistent causal impact of early life exposure to clean water on later life human capital accumulation.[12]

Zhang and Xu (2016) provide an estimate of the impact of access to clean water early in life on educational outcomes by studying the long-term effect of exposure to the rural drinking water program in China.[13] In particular, they use temporal and geographic variation from the first 20 years of implementation of the rural drinking water program and find that individuals exposed to clean water from the ages of 0 to 2 had attained an average of 1.1 additional years of education. The analysis in Zhang and Xu provided a useful first attempt at tackling the question of the persistent impact of access to clean

---

[12] However, Barham (2012) and Venkataramani (2012) provide well-identified evidence linking early life exposure to other infectious diseases to cognitive outcomes. In addition, a number of recent studies record causal impacts of early life infectious disease on education and labor market outcomes for which cognitive development may be a pathway, but they do not explicitly establish impacts on cognition; see Bleakley (2007, 2010), Barreca (2010), Baird,et al (2011), Cutler et al (2010), Lucas (2010). Papers that identify impacts of childhood health shocks other than infectious diseases on cognitive outcomes and achievement include Almond, Edlund, and Palme (2009), Almond and Mazumder (2011), Bharadwaj et al (2012), Maluccio et al (2009), and Stein et al (2005).

[13] Zhang and Xu (2016) also use height as a dependent variable but only provide estimates where effect is measured as an average over all people exposed from age 0 to age 25. Assuming this effect can be used to gauge the early-life access to clean water impact on height is particularly difficult in this case as the authors show that clean water access had a significant and large positive impact on education for individuals that were only exposed for the first time to clean water at ages as old as 21 and 22 years old.

water, but the advantage of our study is that it allows us to improve upon the two important limitations with regard to internal and external validity that exist in their study.

With regard to causal identification, the use of variation in Zhang and Xu that is determined by government provision of resources, which is likely non-random, over a 20 year period raises a concern that bias-inducing selection bias may be present. Moreover, evidence that supports the presence of bias is provided in the paper. Specifically, Zhang and Xu find that individuals exposed for the first time to the rural drinking water program at ages as old as 21 and 22 experience a statistically significant and large, .4 years, increase in their years of education. The fact that the effect for those exposed for the first time at these late age is not statistically distinguishable from the same estimated impact for individuals exposed as early as 3 years old and that in rural China the average years of education is less than 9 years, less than 9% of students in rural China are still in school at age 18, and only 5% of all rural students continue education past high school strongly suggest the estimates may be plagued by nonrandom selection. By studying a program like PAL that was implemented unexpectedly, rapidly, and simultaneously across the entire country of Mexico provides an environment where selection bias is likely limited. Moreover, we can exploit not only timing of exposure through differences in cohorts, but also intensity through heterogeneity in a proxy for pre-intervention water quality at the municipality level. In addition, our estimates will be able to directly control for many forms of potential selection through the inclusion of fixed effects at the municipality, year-month of birth, year-month of interview, and state of birth-year of birth levels, respectively, as well as, uniquely through within family comparisons using maternal fixed effects.

The second limitation faced by Zhang and Xu is that, even if their findings can be interpreted as the causal effect of the rural drinking water program in China, they may not be easily generalizable to the impact of access to clean water outside of China. As the authors note, unlike most developing countries, due to the cultural tradition of eating cooked food and drinking boiled water, the main impurities in the water in China are not disease causing microorganisms. Rather the critical issue with Chinese water is the presence of chemical contaminants from rapid industrialization and country-specific geographical phenomena. Thus, the context and mechanisms under study in Zhang and Xu may not be ideally suited for understanding the environments faced by most policymakers concerned about access to clean water.

The other existing study on this subject in the economics literature comes from Beach et al. (2016). Beach et al. (2016) is an exploration of the impact of early-life exposure to typhoid fever, which is linked to the cleanliness of the water, on educational attainment and earnings during the early 1900s in the United States. They find that eliminating typhoid fever during this time period would lead to a statistically significant one month increase in attained years of schooling. While these results provide some of the only evidence that can plausibly claim to find a causal link between early life exposure to clean water and human capital accumulation, there are two particularly noteworthy distinctions between the contributions of their study and ours.

First, although the authors contend that the primary determinant of typhoid prevalence at this time is the cleanliness of the water and that this is the only mechanism through which the change in educational outcomes can be generated, they do not exploit any specific interventions or public works projects that are known to have generated a change in the water quality. In contrast, the variation we exploit in this paper is generated by the specific implementation of a national clean water program. Moreover, unlike in Beach et al, we explicitly control for the presence of other general health trends that may confound the effect of access to clean water by including childhood mortality rates from diseases, respiratory and vaccine preventable, that vary in the same geographic and temporal way as our estimate for water quality and are indicative of the general health environment, but should be mostly unaffected by PAL.

Second, even if observational variation in local typhoid prevalence is a good and quasi-random proxy for water quality, given that the context of the study is the historical United States the relevance of the significance and magnitude of these estimates to context of the modern developing world is, at best, speculative. The advantage of our study in terms of external validity and policy applicability is that we analyze the impact of access to clean water from a contemporary policy intervention, April 1991, which impacted an entire developing country using nationally representative data.

III. Research Strategy

IIIa. Basic Model

To assess the impact of increased access to clean water during infancy on later life cognition and stature, we estimate:

$$\text{(3)}\quad Y_{ijtmd}=\beta_0+\beta_1 Full_t{*}BaseRate_j+\beta_2 Partial_t{*}BaseRate_j+f(\text{mom's age})_m\gamma+\alpha_1 i.female_i+\eta_{B0}+\omega_t+\lambda_d+\nu_j+Full_t{*}MuniChars_j\psi+Partial_t{*}MuniChars_j\theta+\nu_j{*}YOB_t+\varepsilon_{ijtmd}$$

Y_{imjtd} is the outcome for individual *i*, in municipality *j*, with birth year-month *t*, with mom, *m*, and interviewed in year-month, *d*. $Full_t{*}BaseRate_j$ is an indicator equal to 1 if child, *i*, was exposed to PAL since in utero (i.e. born on or after April 1991), $Full_p$, interacted with the municipality-level gender-specific child diarrheal mortality rate (per 1,000) averaged across the pre-intervention years, 1988-1990, $BaseRate_j$. Thus, the parameter β_1 captures the impact of exposure to increased access to clean-water throughout infancy.[14] We do not observe which individuals are at risk of diarrhea or which benefit from the water reform, so β_1 is interpreted as providing intent to treat (ITT) effects.

Individuals born between 4/1990-3/1991 experienced PAL for part, but not all, of their infancy period. Rather than erroneously adding these partially exposed cohorts to the control or full exposure groups, we estimate their effect separately using $Partial_t{*}BaseRate_j$, where $Partial_t$ is an indicator equal to 1 if the child is born between 4/1990-3/1991 and $BaseRate_j$ is defined as noted previously.

---

[14] We focus on the impact of access to clean water throughout the first year of life because nutritional intake directed towards brain development is highest during this period (Eppig et al 2010). However, we will also use an event study design to test this restriction by estimating impacts at other ages.

By interacting the timing of the intervention with the pre-program municipality-specific diarrhea mortality rates, we exploit municipality*cohort variation in program intensity, as in Acemoglu and Johnson (2007), Bleakley (2007), Bhalotra and Venkataramani (2012). Equation (3) may be thought of as the reduced form of a system in which later life outcomes are allowed to depend upon infant exposure to clean water, with the latter instrumented by the sharp arrival of the National Clean Water program, and whose intensity is allowed to vary by the pre-intervention burden of diarrheal disease in the municipality.

To limit the extent of bias from fixed differences between genders, birth cohorts, interview dates, municipalities, and birth order we add a number of fixed effects to our model. Specifically, the model contains an indicator value for the sex of the respondent, $i.female_i$, and fixed effects for year-month of birth, ω_t, year-month of interview, λ_d, municipality, ν_j, and birth order, η_{BO}. In addition, to account for the potential relationship between the age of the mother and program implementation a quadratic function for the mother's age, $f(\text{mom's age})_m$ is included as well.

Given this set of controls, the primary remaining threats to identification are related to other unobserved trends at the municipality level that are correlated with the variation in our measure of program intensity, $BaseRate_j$. With this issue in mind we add municipality specific linear time trends, n_j*YOB_t, to our model to control for any unobserved linear birth cohort trends that are unique to the respondent's municipality. Despite the use of region-specific linear time trends in addition to the rest of our controls, a recent critique of this type of identification strategy by Jaeger, Joyce, and Kaestner (2018) shows it may not be robust to the inclusion of other pre-intervention characteristic trends. Intuitively, if $BaseRate_j$ is actually capturing variation in a different pre-intervention characteristic, our interpretation of β_1 as being the causal effect of access to clean water may be misguided. In order to account for this we include two control vectors, $Full_t*MuniChars_j$ and $Partial_t*MuniChars_j$ which represent an interaction of the indicators for our cohorts of interest, $Full_t$ and $Partial_t$, with other pre-intervention municipality characteristics such as: percent of population with access to piped water, percent of population with access to sewage system, percent of population has indigenous heritage, mean education level, and mean log of income.

The standard errors of our estimates of equation (3) are clustered at the municipality of birth level to allow for serial correlation in the outcomes across years within municipalities (Bertrand et al 2004). The number of unique municipalities in each analysis ranges between 80-150, which is sufficient to avoid concern over potential over-rejection of the null due to a small number of clusters.

IIIb. Threats to Inference

Given the rich set of controls included in equation (3), there are only two remaining major threats to identification. First, any exploration of the impact of a change in the health environment at early ages must be wary of selective mortality, selection into fertility, and selection into live birth. If certain types of families systematically select into or out of attempts at conception or are more or less likely to be able to have a live birth as a result of increased access to clean water this would generate bias in our estimates. To address the former issue of selection into fertility, we can proceed as is common in this literature and try to limit our exposure to this bias by restricting the sample. In our case we will only include children born in or before 1992. Restricting the sample, though, is only a partial solution to selection into fertility that must rely on a strong assumption of the timing and location of intervention being completely unanticipated. In addition, it does not assist in any way with regard to bias from selection into live birth. Lastly, there is an established relationship between access to clean water and infant mortality (Cutler and Miller, 2005; Field et al., 2011; Galiani et al., 2005; Gamper-Rabindran et al., 2005; Watson, 2005; among others). If, as is likely, the children saved by PAL are non-random, this would generate a bias-inducing selection into our sample. In our analysis we will confront these potential confounders by testing the robustness of our results to limiting our estimates to within-family comparisons.

The second remaining threat to identification in equation (3) is the potential confounding effect of similarly timed public health and education efforts. While there is no record of any government projects of this type[15,16], it is possible that prior interventions could play a role in generating biased estimates of water program impacts if they generate differential pre-trends across the municipalities.[17] To test for bias from these types of unobserved trends we add interactions of our cohorts of interest, $Full_t$ and $Partial_t$, with the pre-intervention levels of two alternative diseases, respiratory disease and vaccine-preventable diseases[18] as controls. These two diseases trends serve as useful controls because infant mortality for these diseases is highly correlated with overall under age 5 mortality, they are driven by several of the same risk factors as diarrheal disease (poverty,

---

[15] Aside from a measles vaccination effort, which is unlikely to be correlated with the temporal and geographic variation of PAL, there were no other public health interventions within a two-year band around the National Clean Water Program (Sepulveda et al 2007).

[16] *Opportunidades* (formerly *Progresa*) is a large and important means-tested cash transfer program in Mexico that provides families with cash conditional upon their children attending schools and health clinics. *Opportunidades* was rolled out across rural Mexican municipalities from 1997, and thus PAL was implemented much earlier, in 1991, and water-treated children, born April 1990 onwards, will have been of school entry age (age 6 or 7) when *Opportunidades* started. Pre-reform birth cohorts in our sample (born 1986-1989), will also have been exposed to *Opportunidades,* but slightly later in their school career. While the age of exposure should prevent any bias from *Opportunidades* on our height results, to generate bias in our estimates on cognition the municipality level intensity of *Opportunidades* exposure would have to be correlated with the municipality level intensity of PAL exposure and the impact of access to *Opportunidades* would have to sharply diverge for kids exposed at 6-7 versus 8-10. While this is unlikely, since *Opportunidades* at this time was restricted to rural areas we can explore whether our results differ by an urban/rural distinction. We find no evidence of this type of relationship across any of our specifications and outcomes (results available upon request).

[17] For example, starting in 1985, the Mexican Federal government introduced and promoted the use of oral rehydration therapy (ORT) for treatment of diarrheal disease (Gutierrez, et al, 1996, Frenk, et al, 2003). If these efforts were targeted to states that performed relatively poorly in terms of infant and child health, some part of the convergence in diarrhea mortality rates observed after 1991 could be attributed to pre-existing trends driven by ORT roll-out.

[18] This category is dominated by measles. As is evident from the blip in vaccine preventable disease in *Figure 2,* there was a measles pandemic in 1989-1990, so including this variable controls for this event. We also re-estimated the equation dropping these years and there was no significant change in the coefficient of interest.

poor nutrition, and crowding), and they are also declining through the sample period. Importantly, though PAL should not meaningfully impact these diseases. Thus, these interactions will capture any municipality-specific trends in the living standards and health environment that are not generated by the change in access to clean water. Specifically, in order to bias our estimates, changes in the disease or general health environment would have to exhibit a trend break in 1991 and be larger in municipalities with higher pre-intervention diarrhea mortality rates. Each of the controls is therefore entered with its pre-1991 level interacted with our cohorts of interest, so as to most severely test the attribution of the 1991 break to the water reform and diarrhea reduction. Robustness of our estimates to these controls implies that omitted trends in common risk factors and interventions that may have created a break in trend in 1991 are unlikely to be biasing our inferences.

Lastly, in an attempt to fully saturate our model and purge our estimates of bias from unobserved regional trends, we assess the robustness of our results to the inclusion of state specific birth year fixed effects. While the amount of identifying variation and power to detect significant differences is reduced when including this set of fixed effects, if we observe little change in the magnitude of the estimates it provides evidence that our conclusions are not being guided by bias from geographic trends.

IV. Data

The individual level data used to estimate equation (3) comes from the Mexican Family Life Survey (*MxFLS*). The MxFLS is a nationally representative survey that collected data from approximately 8,500 households in 150 communities across 16 Mexican states (Rubalcava and Teruel 2007) at baseline in 2002. The MxFLS includes a colored Raven's progressive matrices test and an anthropometric measure of height in addition to household demographics, expenditures, educational attainment, health, labor force participation, and fertility. We limit our analysis to the outcomes for the 1986-1992 birth cohorts who are measured in adolescence, 10-16 years old, during the baseline wave of the MxFLS. Our main analysis will focus on the baseline wave of the *MxFLS* for two reasons.

First, while the overall attrition rate in the MxFLS was quite low across the follow up waves that began in 2005 and 2009, 10% over the entirety of the survey, non-response to certain instruments, including the cognitive tests and anthropometry were considerably higher, approximately 40% and 30%, respectively. Thus, in order to maintain the most representative, and thus externally valid, estimates our primary analysis uses information from the baseline MxFLS survey. That said, we do use the most recent survey wave MxFLS3, which was collected between 2009-2012, to explore the persistence of the impact of access to clean water on cognition and height, as well as, examine other economic outcomes, though we are cautious to examine and document how the loss of sample is likely to bias our results.

The second reason to focus on the baseline MxFLS survey is related to the measure of cognition, colored Raven's progressive matrices test (Raven's). The MxFLS administered a colored Raven's progressive matrices test, which is a non-verbal pattern matching assessment that increases in difficulty as the individual progresses onward, to all individuals above the age of 5.[19] The Raven test is widely used as a test of general intelligence as it is thought to be an informative indicator of an individual's ability to perceive and process information accurately (Raven et al 1984; Stein et al 2005). This test was then repeated in subsequent waves. Thus, a concern with using the later waves of the survey to assess this outcome is that performance on the Raven's test in these waves may reflect learning how to take this specific version of the test rather than cognitive development.

In addition to having systematic measures of cognition and stature, as well as, birth date data[20], the other key component needed to conduct this analysis is locational information that can be used to assign exposure intensity. Using the MxFLS provides two

---

[19] The Raven test is an 18 item test for those aged 5-13 and a 12 item battery for those 14 and over.

[20] The MxFLS provides birth date information at the year-month level for all respondents.

options for assignment of treatment intensity. The first option is to use the individual's municipality of residence during baseline enumeration. This represents the respondent's place of residence when they were between 10 to 16 years old.[21] The advantage of using the individual's municipality of residence is that it is available for all respondents, thus avoiding any loss of sample/power and maintaining the survey's national representativeness. The drawback of this technique is that migration may lead to noise or bias inducing misassignment of exposure intensity. Alternatively, the MxFLS contains some information on the respondent's municipality of birth. Specifically, municipality of birth is recorded for all respondents 15 years old or older in the migration history instrument. Taking responses to this question across the three waves of the survey 78% of the individuals in our analytical sample have a useable municipality of birth. In addition, by matching to the individual's mother's records, an additional 12% can be assigned the mother's place of birth if she has not moved since age 12. [22] For the set of respondents in our sample of interest that have municipality of birth data, we find that it matches the individual's municipality of residence at baseline 89% of the time.

We proceed by generating our main estimates twice, first using municipality of residence at baseline for the entire sample as our assignment location, and then using municipality of birth on the restricted sample with that information as our assignment location. We find that the estimates are qualitatively and quantitatively equivalent. The main difference between the two analyses is a non-trivial reduction in precision when we restrict the sample. Given these results we elect to use the comprehensive municipality of residence data for all subsequent analyses.

Our study primarily focuses on two outcomes. The first is the individual's height-for-age z-score. This measure is calculated using the interviewer-measured height and self-reported age of the respondent and applying the WHO standard growth chart for

---

[21] This is the technique used throughout the early-life exposure to clean water literature (i.e. Zhang and Xu (2016) and Beach et al. (2016)).

[22] The remaining missing municipality of birth information is a result of interviewer notation errors (e.g. spelling or recorded the state of birth instead of the municipality).

developing countries. Our second outcome of interest is the percentile score (0-100) of each respondent on the Raven's colored progressive matrices test. After establishing the relationship between early life access to clean water on these measures of cognitive and physical growth we also explore the impact of PAL exposure in infancy on other human capital outcomes in adolescence such as body mass index, whether the individual still attends school, years of attained education, whether the individual had to repeat any grade levels, and whether the individual is currently employed.

The pre-intervention level of mortality ($BaseRate_j$) by municipality and gender is calculated using the under-5 diarrhea mortality rate over the period 1988-1990. Data on disease mortality rates by cause of death, gender and age were computed using mortality data from the Mexican Secretary of Health (*Secretaría de Salud*) and population estimates from the National Council on Population (*Consejo Nacional de Población*).[23] A one standard deviation in the child diarrheal mortality rate is 10 per 1,000, which is roughly the size of the decline nationally between 1985 and 1995. This metric of change will be used as a guideline for providing relative magnitudes of our estimated effects.

V. The Impact of Early Life Diarrhea Exposure on Human Capital Accumulation

Va. Height and Cognition in Adolescence

Columns 1-3 of Table 2 presents the estimates of models that build up to equation (3) when assigning exposure based on the baseline municipality of residence. Panel A and Panel B provide results when using height-for-age and Raven's test score as the outcome of interest, respectively. Starting in column 1 the model only controls for mother's age, birth order fixed effects, year-month of birth fixed effects, and year-month of interview fixed effects. The estimates in column 1 for both panels suggest that exposure to PAL in

---

[23] The data can be accessed at http://sinais.salud.gob.mx/basesdedatos/ or http://sigsalud.insp.mx/naais/. To construct the indicators for diarrheal and respiratory mortality, we focused primarily on infectious cases. For diarrheal diseases, we used counts for ICD-9 codes A0-A9 and for respiratory diseases, codes 460-466 and 480-487. Vaccine preventable diseases are those from measles, mumps, rubella, diphtheria, and tetanus

infancy was detrimental to physical and cognitive development. This relationship though is likely driven by the fact that areas with worse pre-intervention diarrheal mortality rates, and thus high intensity of exposure, are worse environments for human capital growth in general.

Once fixed differences between municipalities are controlled in column 2, through the inclusion of municipality of birth fixed effects, this adverse relationship completely disappears. In its place we see the first evidence of a non-trivial positive impact of access to clean water in infancy on stature and cognition.

The substantial change in coefficient magnitudes between columns 1 and 2 in Panels A and B are highly suggestive that there are inherent differences between these Mexican municipalities. In an effort to further control these differences Column 3 adds municipality-specific linear time trends and a vector of municipality-specific pre-intervention non-diarrhea characteristic interactions. In both Panels A and B purging the estimates of bias-inducing unobserved trends municipality-specific trends helps to uncover a large positive relationship between access to clean water in infancy and physical and cognitive growth. Specifically, the estimates in column 3 suggest that exposure to a one standard deviation decrease in the diarrheal mortality rate (i.e. 10 per 1,000 decline) from PAL in infancy led to .11 s.d. increase in height and a 1.7 percentile point increase in Raven's test score.

While the results in column 3 provide evidence of the persistent beneficial impact of access to clean water in infancy, there is potential that these results are biased as a result of mortality selection, fertility selection, or selection into live birth resulting from PAL. Thus, in column 4 we introduce maternal fixed effects into equation (3).[24]

---

[24] When employing a within-family analysis it is important to confirm that the compositional change in the sample is leading to result that are not generalizable or comparable to the full sample. In our case, this does not appear to be a concern as results analogous to those found in columns 1-3 of Panels A and B but using the mother-fixed effects sample are qualitatively similar and, if anything, suggest the within-family comparisons may be providing a lower bound. These estimates are found in Appendix Table 1.

The introduction of within-family comparisons in column 4 only strengthens the evidence that exposure to PAL in infancy had a long-lasting impact on physical and cognitive development. In Panel A, the effect of access to clean water in infancy is equal in size to what was found in column 3, while in Panel B, the removal of bias related to selective parentage has strengthened and confirmed the relationship between early life exposure to clean water and cognition.

The results in column 4 provide an indication that there is some PAL related negative selection into the sample. We directly explore this by re-estimating equation (3) but replacing the outcome of interest with maternal characteristics. This analysis, provided in Appendix Table 2, shows that the sample of individual's exposed to greater PAL intensity in infancy are more likely to have less intelligent, poorer, and less healthy mothers, as indicated by the mother's Raven's test score, household per capita expenditure and the mother's likelihood of being obese (BMI>30) and underweight (BMI<18.5), respectively.[25]

There are three primary reasons this change in composition exists. The most likely cause of this selection is the well-documented relationship between access to clean water and the reduction of infant mortality generally (Cutler and Miller, 2005; Field et al., 2011; Galiani et al., 2005; Gamper-Rabindran et al., 2005; Watson, 2005; among others) and in the specific context of PAL (Bhalotra et al., 2017). Given that the increase in survival is likely occurring at the bottom of the health distribution, this change would likely lead to an increase in the number of children in the survey from poorer or less healthy households in places that had the most intense exposure to PAL. Additionally, if we think PAL is lowering the cost of child quality, it could induce less well off families who were

---

[25] Children exposed to a 1 standard deviation increase in PAL intensity on average have mothers with 21% less household per capita expenditure and that are almost 3 times more likely to be underweight (i.e. BMI<18.5). The difference in mother's Raven's score and likelihood obesity while large in magnitude, an average decline of 3% for cognition and an average 8% increase for obesity for a 1 standard deviation increase in PAL intensity, are only marginally significant at the .13 and .15 level, respectively.

previously on the margin of family expansion due to budget constraints to select into fertility. Lastly, access to clean water may also cause selection into live birth. If the positive health shock of PAL creates a healthier environment for pregnant women, some pregnancies that previously were not viable will now conclude in a live birth. As with infant mortality, this selection would be the most pronounced at the bottom of the family wealth and health distribution and lead to the type of selection found in Appendix Table 2. Regardless of the exact cause(s), removing this negative bias inducing selection with the within-family comparisons is what allows the considerably larger estimate of PAL exposure's impact on cognition to come through in column 4 of Panel B as compared to column 3.

Finally, columns 5 and 6 complete our analysis by additionally including controls for municipality-specific trends in the general health environment and state of birth-birth year specific unobserved factors, respectively. In each case, for both outcomes, the introduction of these controls either confirms or increases the positive relationship between access to clean water in infancy and persistent gains in cognitive and physical growth.[26] Specifically, the estimates in column 6 suggest that experiencing a 1 s.d. intensity of the PAL effect (i.e. a 10 in 1,000 decline in the under 5 diarrheal mortality effect) in infancy will on average lead to a .15 standard deviation increase in both height and cognitive assessment score (3.4 percentage point increase relative to a standard deviation of 22.25) in adolescence.

As mentioned in section IV, one potential concern with interpreting and generalizing the estimates in Table 2 is that in order to utilize the entire analytical sample we assign treatment intensity using municipality of residence in adolescence rather than

---

[26] Since the addition of the alternative disease trend controls and state-year fixed effects does not substantially alter the estimates but potentially cause a reduction in efficiency, we do not include them when reporting within-family estimates of the other outcomes explored in the paper. As in Table 2, the inclusion or exclusion of these controls does not significantly change the magnitude of the estimates or conclusions drawn from the analysis.

the more limited information we have available on each respondent's municipality of birth. In order to assess whether this choice is leading to misinterpretation of the impact of access to clean water in infancy, we re-estimate equation (3) using municipality of birth responses to assign exposure on the restricted sample with that information. These results are provided in Table 3 and the magnitudes of the coefficient estimates closely mirror and not statistically distinguishable from those from in Table 2. The critical difference between the findings in Table 2 and Table 3 is the non-trivial precision gains from using the full sample. Since this is achieved without generating bias in the estimates and allows us to avoid the potential external and/or internal validity concerns of attrition we proceed in the rest of the analyses to use municipality of residence in adolescence as our location of treatment assignment.

Vb. Event Study Analysis

Two important questions remain when assessing the results in Table 2. First, is there any evidence of an increasing trend in human capital outcomes in areas with higher levels of pre-PAL diarrheal mortality. While we have shown that including municipality-specific linear trends, vectors of pre-intervention municipality of birth characteristics and alternative mortality rates interacted with our cohorts of interest, as well as, state-year fixed effects in equation (3) does not change our conclusions, there is still the possibility that we have not captured all possible confounding unobserved trends.

Second, our estimates thus far have focused on the impact of access to clean water throughout infancy. This choice is motivated by the epidemiological and biological evidence that both the incidence of diarrhea (marking treatability by the water reform or treatment intensity)[27] and the rate of physical and cognitive development (marking

---

[27] The *MxFLS* survey queries the incidence of diarrhea in the two weeks preceding the date of interview. As many as 25% of infants in the sample had diarrhea in the two weeks preceding the survey, compared with 5 to 10% for 5 to 50 year olds.

elasticity of the outcome to the treatment) are greatest during this period. Empirically though, we can directly explore the impact at older ages of first exposure as well.

In order to provide evidence related to both of these concerns we estimate an event study analysis. Specifically, we re-estimate the within-family difference version of equation (3) for the 1986-1992 birth cohorts, but replace $Full_t$ and $Partial_t$ with a vector of indicator variables for every birth cohort in the sample except 1989.[28] As is standard, our exclusion cohort is chosen as the year just prior to exposure. In our case, 1989, is the birth cohort just prior to our first exposure group, $Partial_t$. Since in our data cohort and period coincide, the results in this sub-section also implicitly confirm the absence of trend breaks in the coefficient series at dates other than 1991. Thus the event study acts as a placebo test of whether we have identified impacts flowing from the water reform rather than some other confounding trend.

The results of the event study analyses are provided in Figures 6 and 7 for height-for-age and Raven's score, respectively. In both cases, the coefficients for birth cohorts exposed for the first time after infancy are small and insignificant. Moreover, there is a noticeable jump in the event study coefficients for both the impact on stature and cognition for the cohorts exposed during infancy. As a whole the event study analyses provide evidence supportive of the assumptions that our estimates are not driven by unaccounted for trends correlated with the timing and intensity of PAL and that the most advantageous exposure period is in infancy.

Vc. Persistence of the PAL Effect on Height and Cognition into Early Adulthood

The most rich and representative data that is available to explore the impact of PAL exposure in infancy on cognitive and physical development were recorded in the baseline MxFLS survey when the cohorts of interest were in adolescence. An additional

---

[28] In the event study analysis all birth cohorts are redefined as containing those born from 4/1/YOB-3/31/YOB+1. This allows us to more closely match our relevant cohorts from the main analysis, $Full_t$ and $Partial_t$. Thus for example, in the event study the 1991 cohort is defined as those born between 4/1/1991-3/31/1992.

advantage of the MxFLS that we can exploit for our study is the ability to see if the effects on height and intelligence we uncovered in adolescence at baseline persist into early adulthood (17-26 years of age) in MxFLS3. Being able to examine the persistence of gains produced by an early-life health shock by studying the same cohorts at two different points in time is a novel and unique opportunity in this literature. The complication that arises with this type of analysis, and one of the reasons we employ the baseline survey for out main results, is that by MxFLS3 we no longer have outcome information for all of our respondents. In fact, while overall survey attrition between the baseline MxFLS survey and MxFLS3 was only slightly over 10%, non-response to Raven's and height measurement was considerably higher. Amongst the sample of interest we are missing MxFLS3 height records for ~35% and Raven's score information for ~45% of our respondents.

In order to assess the way this will impact our interpretation of findings using the available MxFLS3 sample we start by comparing our main estimates in Table 2 to analogous estimates using MxFLS baseline data but only on the sample with available MxFLS3 data. Column 1 of Table 4 reproduces results from our main analysis in Table 2 and then column 2 estimates this same model but restricted only to the respondents with MxFLS3 data for the outcome of interest. In panels A and B we see a sizable decrease in the estimated effect size for both height-for-age and cognition. This suggests that non-response for our outcomes of interest in MxFLS3 is systematically higher for respondents that experienced the largest benefits to physical and cognitive development from PAL exposure in infancy. Thus, we would expect that our analysis of the relationship between access to clean water in infancy and height and intelligence in early adulthood using the MxFLS3 provide a lower bound estimate.

After establishing the bias that will be generated by moving to the MxFLS3 sample, we next re-estimate our models using data from the MxFLS3 survey in columns 3-5. In Panel A, despite using a sample we know will cause our estimates to be biased towards 0, in our preferred specifications in columns 4-6 that use within-family variation and thus can mitigate some non-response selection-bias, the average of our coefficients indicates a positive impact on height for age in early adulthood of .14 standard deviations.[29] This magnitude matches the analogous average coefficient estimate in our main analysis in Panel A of Table 2.[30]

Similarly, while we believe the MxFLS3 sample will lead to an attenuate estimate of the impact of PAL exposure in infancy on cognition, the results in Panel B of Table 4 still provide evidence of a persistent effect of the program. As with height, once some of the confounding impacts of the changed sample composition are accounted for with maternal fixed effects in columns 4-6, we find coefficient magnitudes that are similar or larger in size to the one's found in Panel B of Table 2. Our estimates suggest that, similar to our baseline results, experiencing a one standard deviation improvement in diarrheal mortality rates in the respondent's municipality of birth from PAL in infancy leads to an 4-10% (2.5-5.4 points relative to an average of 55.5) increase in cognitive tests scores on average.[31,32] Overall, while the loss in power due to non-response has reduced the precision of our estimates using MxFLS3 data and likely generated bias towards 0, the results in Table 4 provide a strong indication that the gains in stature and cognition from exposure in infancy to PAL are long-lasting.

---

[29] Estimating the impact of PAL exposure on height in early adulthood when assigning exposure using municipality of birth information provides a similar effect size of a .13 standard deviation increase in height using our preferred within-family specification.

[30] While in MxFLS1 the age range of our cohorts allows us to use WHO child growth charts to estimate height-for-age, that is not the case in MxFLS3. Thus, in MxFLS3 we use simple height for age z-scores that are calculated as (height-mean height)/standard deviation of height separately for each age, where the mean and standard deviation are calculated for our cohorts of interest, the 1986-1992 birth cohorts.

[31] The estimates in Panel B of Table 4 are similar or larger than those in Panel B of Table 2, but the significant (45%) loss in sample size makes them much less precisely estimated.

[32] Estimating the analogous impact of PAL exposure on cognition in early adulthood when assigning exposure using municipality of birth information provides a similar effect size of a 3-8% increase.

Vd. Other Human Capital and Economic Outcomes in Adolescence

In addition to examining the impact of early life access to clean water on direct measures of cognitive and physical growth, the richness of the MxFLS allows to also explore if exposure to PAL in infancy led to direct changes in their educational outcomes and economic behavior. We start by examining the impact of increased access to clean water in infancy on attained education in adolescence by estimating the within-family version of equation (3).[33] This analysis is found in column 1 of Table 5 and provides evidence of gains to attained education for children that gained access to clean water throughout infancy. The estimate suggests that a one standard deviation decrease in the under-5 diarrheal mortality rate throughout infancy led to a statistically significant 4% increase in years of attained education (.2 years compared to an average of 5.3).

This change in educational attainment is likely the result of one of two channels, either the kids exposed to PAL during infancy are staying in school longer or are progressing in school faster. To parse these two mechanisms we proceed in column 2 to test if PAL exposure led to changes in likelihood that our respondents are currently attending school. There is no evidence of any impact on this behavior. In order to explore school progression, in columns 3 and 4, we estimate our within-family version of equation (3) using whether the individuals has repeated any grade and whether the individual has started secondary school as our outcomes of interest, respectively. In both cases there are large and statistically significant differences for individuals with greater PAL exposure throughout infancy, as are found to repeat fewer grades and to be more likely to have progressed to secondary school.

Lastly, since employment even at these young ages is not uncommon in Mexico (12.5% of our sample of 10-16 year olds) we examine if the decision to join the labor market as early as adolescence is affected by access to clean water early in life. Column 5 of Table 5 provides our findings when using whether the individual is currently employed

---

[33] As in Table 2, results are qualitatively and quantitatively similar if the alternative disease trends and state-year of birth fixed effects are also included in the model.

as our outcome of interest, which indicate that PAL exposure in infancy is leading to a delay in labor market engagement. It is possible that this decision reflects a tradeoff between current income and the long-term gains from greater investment in education that those with greater cognitive and physical capacity are willing to make.

Ve. Other Human Capital and Economic Outcomes in Early Adulthood

As with our exploration of the impact of access to clean water on cognitive and physical development into early adulthood, the longitudinal nature of the MxFLS also allows to also investigate how changes in educational and economic outcomes persist later in life. As before we rely on data from the MxFLS3 to provide this analysis, but unlike when examining stature and intelligence, the loss of sample due to non-response is much less stark. Specifically, the loss in our sample of interest from MxFLS1 to MxFLS3 on education and employment outcomes is only 4 and 20%, respectively.

In column 1 of Table 6 we explore whether the boost in educational attainment we found in adolescence continues into early adulthood. While the estimated effect remains positive, it is considerably smaller in size and no longer statistically distinguishable from 0. Given our previous findings that the primary mechanism behind the difference in years of education in adolescence was not a result of school dropout but rather speed of progression, the findings in Table 6 suggest eventually the students terminate their studies at an equivalent level.

While the gains in cognition and stature do not lead to different educational attainment, we next examine if it change an individual's labor market engagement. In column 2 of Table 6 we see that there is no evidence that early life exposure to PAL changes an individual's likelihood of being employed in early adulthood.

With about half of our sample of interest actively working in early adulthood, the better early life health environment for the PAL exposed individuals does not seem to impact whether they are employed or not. Despite no difference in labor market participation, the benefit to the cognitive and physical advantage possessed by those who had access to clean water during infancy may show up in their level of productivity, either through greater salary or reduced hours of work. In columns 3-4 we investigate these economic channels for those in the labor market by estimating our within-family version of equation (3) on the outcomes of total hours worked last week and earning per hour, respectively.[34] The results found in columns 3-5 tell a compelling story of increased productivity for individuals with greater access to clean water throughout infancy. While estimated on a small sample, the findings suggest those exposed to a 1 standard deviation in decrease in the under-5 diarrhea mortality rate due to PAL throughout infancy are earning 24 pesos more per hour, which represents a 100% increase compared to the average. Thus, while access to clean water has not led to more education or likelihood of employment, it has, through an increase in physical and cognitive growth, delivered the opportunity for much higher paying positions.

Vf. Gender Differences in Program Effects

It is common in both the experimental and quasi-experimental literature on the relationship between an improved health environment in childhood and long-term human capital accumulation to find that the impacts differ significantly with respect to the gender of the child (Almond and Mazumder, 2011; Banerjee et al., 2010; Bhalotra and Venkataramani, 2012; Cutler et al., 2010; Macinni and Yang, 2009; Miguel and Kremer, 2004; Pitt, Rosenzweig, and Hassan 2012; Rosenzweig and Zhang, 2013; among others). In order to explore the presence of gendered effects of PAL exposure in infancy we re-estimate our main analysis but additionally interact both $Full_t*BaseRate_j$ and $Partial_t*BaseRate_j$ with an indicator for the respondent being male, as well as, our continuous control variables and pre-municipality characteristic trends.[35]

---

[34] There is no direct earning per hour measure in the MxFLS, so this is estimated as earnings per month divided by hours worked last week times four.

[35] The estimates are not sensitive to additionally allowing cohort, interview date, and/or birth order effects also to vary by gender.

The estimates for the direct effect of PAL exposure in infancy (i.e. the impact on girls), the interaction effect (i.e. the difference in impact between boys and girls), as well as, the p-value for the sum of these two effects (i.e. the impact on boys) are provided in Table 7.[36] In Panel A, columns 1-3 suggest that the benefits to height from access to clean water are more pronounced for women, but the introduction of gender specific pre-intervention municipality characteristic trends greatly alters that interpretation. Once differences in these trends is accounted for it is clear that the gains to stature accrued through early life PAL exposure are gender neutral.

Panel B, on the other hand, does provide evidence of a gendered impact on cognition. Exposure to PAL in infancy for girls leads to a non-trivial and significantly larger gain to cognition in adolescence as compared to boys. Specifically, a one standard deviation improvement in diarrheal mortality rates in the respondent's municipality of birth from PAL in infancy generates an 8.7% increase in cognitive test scores for girls and only a 3.3% increase for boys. It is important to note, that while imprecisely estimated and significantly smaller than the effect for girls, the magnitude for the impact on cognition for boys should not be considered economically insignificant. For instance, similarly sized estimates when using the full sample size, as in Column 2 of Table 2, are statistically different from 0. In support of the idea that males are also experiencing important cognitive benefits to early life exposure to clean water, Table 8 shows that the increased educational progress effects and delay in labor market engagement found in Table 6 are predominately a change experienced by males. Finally in Table 9 we explore if the boost in productivity in early adulthood from infant exposure to clean water, previously shown in Table 7, is gender specific and find no evidence that this is the case.

---

[36] The gender specific partial impact for boys and girls born between April 1990 and March 1991 are estimated in these models but not reported in Table 5.

VI. Potential Alternative Mechanisms

We interpret our estimates as stemming from access to clean water in infancy that decreased the likelihood of diarrhea infections that would have directly impair physical and cognitive development in a critical period by diminishing net nutrition.[37] Here we consider the plausibility of competing explanations. First, we consider whether the primary mechanism may be a reduction in the *adult* burden of disease which impacts children's physical and cognitive development either through productivity-led improvements in household income or through improved maternal health leading to better fetal and infant health. This is unlikely given the low levels of adult diarrhea incidence. It is estimated that diarrhea morbidity rates in the developing world for adults are at most 0.3 to 0.8 episodes per year in contrast with around 3 episodes per year for children under the age of five (Fischer Walker and Black 2010, Fischer Walker et al 2012). Also, if the channel were income related, we may expect to see impacts through childhood,[38] but our finding that effects fade after infancy and are hardly discernible after the age of one is consistent with what is initially a biological mechanism.

VII. Conclusions

A large body of work highlights the importance of strength and cognitive skills in driving socioeconomic outcomes, making it important from a public policy standpoint to understand how to narrow gaps in physical and cognitive development. We find that access to clean water in early-life, through its reduction of infectious disease exposure, has a positive causal influence on cognition and stature. Specifically, our estimates suggest that that exposure in infancy to a one standard deviation decrease in the diarrheal mortality rate (i.e. 10 per 1,000 decline) as a result of a clean water intervention led to a .15 standard deviation increase in both height and cognitive assessment score.

---

[37] The positive shock of access to clean water in infancy may also be reinforced through the increased presence and productivity of subsequent parental investments.

[38] For instance, Dahl and Lochner (2012) examine income shocks and record impacts for children above the age of 2.

These magnitudes compare favorably to those from other health (or nutritional) interventions/shocks. In particular, these gains are similar to those estimated in Bharadwaj et al (2013), who find that more intensive care for low birth weight neonates resulted in 0.1 and 0.2 standard deviation increases in language and mathematics scores, respectively, in a sample of 10-16 year olds. The costs of water reform are likely much smaller than the costs of advanced neonatal care. Almond, Mazumder and Van Ewijk (2011) show that Ramadan fasting among pregnant mothers led to 0.05 to 0.08 standard deviation decrease in test scores for 7-year old Muslim children in British schools. Birth year exposure to malaria eradication in 1950s Mexico led to a 0.1-0.2 standard deviation increase in adult Raven scores for men (Venkataramani 2012) and increased access to an early childhood and family planning program in Bangladesh led to a 0.39 standard deviation increase in mini-Mental State exam scores (Barham 2012).

In addition, we provide evidence to show that these gains in physical and cognitive growth persist into early adulthood and lead to significant increases in labor market productivity. These results imply that clean water provision, acting through infant disease exposure, can potentially help induce economic growth and shrink the classic cross-country differences between developed and developing countries.

References

Abramitzky, Ram and Victor Lavy. 2012. "How Responsive is Investment in Schooling to Changes in Redistributive Policies and in Returns." *Mimeo*, Stanford University.

Acemoglu, Daron and Simon Johnson. 2007. "Disease and Development: The Effect of Life Expectancy on Economic Growth." *Journal of Political Economy*, 115(6), 925-65.

Adhvaryu, Achyuta and Anant Nyshadham. 2012. "Endowments at Birth and Parents' Investments in Children." *Mimeo, Yale University.*

Aguayo-Tellez, Ernesto; Jim Airola and Chinhui Juhn. 2010. "Did Trade Liberalization Help Women? The Case of Mexico in the 1990s." *NBER Working Paper w16195.*

Ahuja, Amrita; Michael Kremer and Alex P Zwane. 2010. "Providing Safe Water: Evidence from Randomized Evaluations." *Annual Review of Resource Economics*, 2, 237-56.

Aizer, Anna and Flavio Cunha. 2012. "The Production of Human Capital in Childhood: Endowments, Investments and Fertility." *Mimeo, Brown University.*

Almond, Douglas; Lena Edlund and Marten Palme. 2009. "Chernobyl's Sublcinical Legacy: Prenatal Exposure to Radioactive Fallour and School Outcomes in Sweden." *Quarterly Journal of Economics*, 124(4), 1729-72.

Almond, Douglas and Bhashkar Mazumder. Forthcoming. "Fetal Origins and Parental Responses." *Annual Reviews of Economics.*

Almond, Douglas and Bhaskar Mazumder. 2011. "Health Capital and the Prenatal Environment: The Effect of Ranamdan Observance During Pregnancy." *American Economic Journal: Applied Economics*, 3(4), 56-85.

Almond, Douglas and Janet Currie. 2011. Human Capital Development Before Age Five. *Handbook of Labor Economics*, Volume 4b, Elsevier, North Holland.

Apps, Patricia; Silvia Mendolia and Ian Walker. 2012. "The Impact of Pre-School on Adolescents' Outcomes: Evidence from a Recent English Cohort." *IZA Discussion Paper No. 6971.*

Ashraf, Sania; M Hamidal Huque; Eben Kenah; Mubina Agboatwalla and Stephen P. Luby. 2013. "Effect of Recent Diarrheal Episodes on Risk of Pneumonia in Children

Under the Age of 5 Years in Karachi, Pakistan." *International Journal of Epidemiology*, 42(1), 194-200.

Baird, Sarah; Joan Hamory Hicks; Michael Kremer and Edward Miguel. 2011. "Worms and Work: Long-Run Impacts of Child Health Gains." *mimeo, University of California, Berkeley.*

Barham, Tania. 2012. "Enhancing Cognitive Function: Medium-Term Effects of a Health and Family Planning Program in Matlab." *American Economic Journal: Applied Economics*, 4(1), 245-73.

Barham, Tania; Karen Macours, and John A. Maluccio. 2013. "Boys' Cognitive Skill Formation and Physical Growth: Long-Term Experimental Evidence on Critical Ages for Early Childhood Interventions". *American Economic Review: Papers & Proceedings*, 103(3): 467–471.

Behrman, Jere; Yingmei Cheng and Petra Todd. 2004. "Evaluating Preschool Programs When Length of Exposure to the Program Varies: A Nonparametric Approach." *Review of Economics and Statistics*, 86(1), 108-32.

Behrman, Jere; Andrew Foster; Mark Rosenzweig and Prem Vashishtha. 1999. "Women's Schooling, Home Taching, and Economic Growth." *Journal of Political Economy*, 107(4), 682-714.

Behrman, Jere; Susan Parker and Petra Todd. 2009. "Medium-Term Impacts of the American Oportunidades Conditional Cash Transfer Program on Rural Youth in Mexico," S. Klasen and F. Nowak-Lehmann, *Poverty, Inequality, and Policy in Latin America.* Cambridge: MIT Press,

Bertrand, M; E Dulfo and S Mullainathan. 2004. "How Much Should We Trust Differences-in-Differences Estimates?" *Quarterly Journal of Economics*, 119, 249-75.

Bhalotra, Sonia; David Hollywood and Atheendar Venkataramani. 2012. "Fertility Responses to Reductions in Infant and Maternal Mortality: Quasi-Experimental Evidence from 20th Century America." *Mimeo, University of Bristol.*

Bhalotra, Sonia and Atheendar Venkataramani. 2012. "Shadows of the Captain of the Men of Death: Early Life Health Interventions, Human Capital Investments, and Institutions." *Mimeo, University of Bristol.*

Bharadwaj, Prashant; Giacomo De Giorgi; David Hansen and Christopher Neilson. 2012. "The Gender Gap in Mathematics: Evidence from Low- and Middle-Income Countries. ." *Mimeo, University of California, San Diego.*

Bharadwaj, Prashant; Katrine Vellesen Loken and Christopher Neilson. 2013. "Early Life Health Interventions and Academic Achievement." *American Economic Review*, 103(5), 1862-91.

Bhutta, Zulfiqar A; Jai K Das; Neff Walker; Arjumand Rizvi; Harry Cambell; Igor Rudan and Robert E Black. 2013. "Interventions to Addres Deaths from Childhood Pneumonia and Diarrhoea Equitably: What Works and at What Cost? " *The Lancet*, 381(9875), 1417-29.

Bleakley, Hoyt. 2007. "Disease and Development: Evidence from Hookworm Eradication in the American South." *The Quarterly Journal of Economics*, 122(1), 73-117.

Bleakley, Hoyt. 2010. "Health, Human Capital and Development." *Annual Review of Economics*, 2, 283-310.

Bozzoli, C; A Deaton and C Quintana-Domeque. 2009. "Adult Height and Childhood Disease." *Demography*, 46, 647-69.

Cameron, A Colin; Jonah Gelbach and Douglas Miller. 2008. "Bootstrap-Based Improvements for Inference with Clustered Errors." *Review of Economics and Statistics*, 90, 414-27.

Case, Anne; Christina Paxson. 2008. "Stature and Status: Height, Ability, and Labor Market Outcomes." *Journal of Political Economy*, 116(3), 499-532.

Chay, Kenneth Y; Jonathan Guryan and Bhashkar Mazumder. 2009. "Birth Cohort and the Black-White Achievement Gap: The Roles of Access and Health Soon after Birth." *NBER Working Paper w15078.*

Costa, Dora. 2013. "Health and the Economy in the United States, from 1750 to the Present," NBER Working Papers 19685, National Bureau of Economic Research, Inc.

Cunha, Flavio and James Heckman. 2007. "The Technology of Skill Formation." *American Economic Review*, 97(2), 31-47.

Cunha, Flavio; James Heckman and Susan Schennach. 2010. "Estimating the Technology of Cognitive and Noncognitive Skill Formation." *Econometrica*, 78(3), 883-931.

Cutler, David; Winnie Fung; Michael Kremer; Monica Singhal and Tom Vogl. 2010. "Early-Life Malaria Exposure and Adult Outcomes: Evidence from Malaria Eradication in India." *American Economic Journal: Applied Economics*, 2, 72-94.

Cutler, David; Felicia Knaul; Rafael Lozano; Oscar Mendez and Beatriz Zurita. 2002. "Financial Crisis, Health Outcomes and Ageing: Mexico in the 1980s and 1990s." *Journal of Public Economics*, 84(2), 279-303.

Cutler, David and Grant Miller. 2005. "The Role of Public Health Improvements in Health Advances: The Twentieth-Century United States." *Demography*, 42(1), 1-22.

Dahl, Gordon B and Lance Lochner. 2012. "The Impact of Family Income on Child Achievement: Evidence from Changes in the Earned Income Tax Credit." *American Economic Review*, 102(5), 1927-56.

Deverman, Benjamin E and Paul H Patterson. 2009. "Cytokines and Cns Development." *Neuron*, 64, 61-78.

Economist, The. 2010. Disease and intelligence: Mens sana in corpore sano. See http://www.economist.com/node/16479286

Eppig, Christopher; Corey L Fincher and Randy Thornhill. 2010. "Parasite Prevalence and the Worldwide Distribution of Cognitive Ability." *Proceedings of the Royal Society B*, 277(1701), 3801-8.

Fernald, Lia; Paul Gertler and Lynnette Neufeld. 2009. "10-Year Effect of Oportunidades, Mexico's Conditonal Cash Transfer Programme, on Child Growth, Cognition, Language, and Behavior: A Longitudinal Follow-up Study." *The Lancet*, 374(9706), 1997-2005.

Fischer Walker, Christa and Robert Black. 2010. "Diarrhoea Morbidity and Mortality in Older Children, Adolescents, and Adults." *Epidemiology and Infection* 138(9), 1215-26.

Fischer Walker, Christa L; Igot Rudan; Li Liu; Harish Nair; Evropi Theodoratou; Zulfiqar Bhutta; Katherine L O'Brien; Harry Campbell and Robert E Black. 2013. "Global Burden of Childhood Diarrhoea and Pneumonia." *The Lancet*, 381(9875), 1405-16.

Fischer Walker, Christa; Laura Lamberti; Linda Adair; Richard Guerrant; Andres Lescano; Reynaldo Martorell; Relena Pinkerton and Robert E Black. 2012a. "Does Childhood Diarrhea Influence Cognition Beyond the Diarrhea-Stunting Pathway?" *PLoS One*, 7(10), e47908.

Fischer Walker, Christa; Jamie Perin; Martin Aryee; Cynthia Boschi-Pinto and Robert Black. 2012b. "Diarrhea Incidence in Low- and Middle-Income Countries in 1990 and 2010: A Systematic Review." *BMC Public Health*, 12(220), doi: 10.1186/471-2458-12-220.

Frenk, J; J Sepulveda; O Gomez-Dantes and F Knaul. 2003. "Evidence-Based Health Policy: Three Generations of Reform in Mexico." *Lancet*, 362(9396), 1667-71.

Fryer, Roland and Steven Levitt. 2010. "An Empirical Analysis of the Gender Gap in Mathematics." *American Economic Journal: Applied Economics*, 2, 210-40.

Galiani, Sebastian; Paul Gertler and Ernesto Schargrodsky. 2005. "Water for Life: The Impact of the Privatization of Water Services on Child Mortality." *Journal of Political Economy*, 113(1), 83-120.

Galor, Oded and David Weil. 1996. "The Gender Gap, Fertility, and Growth." *American Economic Review*, 86(3), 374-87.

Gamper-Rabindran, Shanti; Shakeeb Khan and Christopher Timmins. 2010. "The Impact of Piped Water Provision on Infant Mortality in Brazil: A Quantile Panel Data Approach." *Journal of Development Economics*, 92(2), 188-200.

German-Soto, Vincente. 2005. "Generación Del Producto Interno Bruto Mexicano Por Entidad Federativa, 1940-1992,." *El Trimestre Económico*, 72, 617-53.

Grantham-McGregor SM, Cheung YB, Cueto S, Glewwe P, Richter L, Strupp B and International Child Development Steering Group (Engle P, Black M, Meeks Gardner J, Lozoff B, Wachs T, Walker S). 2007. "Child development in developing countries 1. Developmental potential in the first five years for children in developing countries. " *Lancet* 369: 60-70

Guerrant, David; Sean Moore; Aldo Lima; Peter Patrick; John Schorling and Richard Guerrant. 1999. "Association of Early Childhood Diarrhea and Cryptosporidiosis with Impaired Physical Fitness and Cognitive Function Four-Seven Years Later in a Poor Urban Community in Northeast Brazil." *American Journal of Tropical Medicine and Hygiene*, 61(5), 707-13.

Günther, Christian M., Alexander Bürger, Markus Rickert, Alexander Crispin, and Christoph U. Schulz. 2008, "Grip Strength in Healthy Caucasian Adults: Reference Values," *Journal of Hand Surgery* 33: 558-565.

Gutierrez, R; R Tapia-Conyer; H Guiscafre; H Reyes; H Martinez and J Kumate. 1996. "Impact of Oral Rehydration and Selected Public Health Interventions on Reduction of Mortality from Childhood Diarrhoel Diseases in Mexico." *Bulletin of the World Health Organization*, 74(2), 189-97.

Hanson, Gordon. 2003. "What Has Happened to Wages in Mexico since Nafta? Implications for Hemispheric Free Trade." *NBER Working Paper w9563*.

Heckman, James; Jora Stixrud and Sergio Urzua. 2006. "The Effect of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior." *Journal of Labor Economics*, 24, 411-82.

Instituto Mexico de Technologia del Agua. 2008. "Evaluacion De Consistencia Y Resultados 2007," C. N. d. Agua, Mexico City, Mexico:

Jayachandran, Seema; Adriana Lleras-Muney and Kimberly V Smith. 2010. "Modern Medicine and the 20th-Century Decline in Mortality: Evidence on the Impact of Sulfa Drugs." *American Economic Journal: Applied Economics*, 2(2), 118-46.

Jensen, Robert. 2010. "The (Perceived) Returns to Education and the Demand for Schooling." *Quarterly Journal of Economics*, 125(2), 515-548.

Jensen, Robert. 2012. "Do Labor Market Opportunities Affect Young Women's Work and Family Decisions? Experimental Evidence from India." *Quarterly Journal of Economics*, 127(2), 753-792.

Kraemer, Sebastian. 2000. "The Fragile Male." *British Medical Journal*, 321(7276), 1609-12.

Lucas, Adrienne. 2010. "Malaria Eradication and Educational Attainment: Evidence from Paraguay and Sri Lanka." *American Economic Journal: Applied Economics*, 2, 46-71.

Guiso, Luigi; Fernando Monte, Paola Sapienza and Luigi Zingales "Culture, Gender and Math", Science, 2008

Macours, Karen; Norbert Schady and Renos Vakis. 2012. "Cash Transfers, Behavioral Changes, and Cognitive Development in Early Childhood: Evidence from a Randomized Experiment." *American Economic Journal: Applied Economics*, 4(2), 247-73.

Maluccio, John; John Hoddinott; Jere Behrman; Reynaldo Martorell; Agnes Quisumbing and Aryeh Stein. 2009. "The Impact of Improvement Nutrition During Early Childhood on Education among Guatemalan Adults." *The Economic Journal*, 119(537).

Mathiowetz, Virgil, Nancy Kashman, Gloria Volland, Karen Weber, Mary Dowe, and Sandra Rogers. 1985, "Grip and Pinch Strength: Normative Data for Adults," *Archives of Physical Medicine and Rehabilitation* 66: 69-72.

Miguel, Edward and Michael Kremer. 2004. "Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities." *Econometrica*, 72(159-217).

Niehaus, Mark; Sean Moore; Peter Patrick; Lori Derr; Breyette Lorntz; Aldo A Lima and Richard Guerrant. 2002. "Early Childhood Diarrhea Is Associated with Diminished Cognitive Function 4 to 7 Years Later in Children in a Northeast Brazilian Shantytown." *American Journal of Tropical Medicine and Hygiene*, 66(5), 590-93.

Persico, Nicola; **Andrew Postlewaite; Dan Silverman.** 2004. "The Effect of Adolescent Experience on Labor Market Outcomes: The Case of Height" ***Journal of Political Economy,*** 112(5), 1019-1053.

Pitt, Mark M; Mark Rosenzweig and Mohammed Nazmul Hassan. 2012. "Human Capital Investment and the Gender Devision of Labor in a Brawn-Based Economy." *American Economic Review*, 102(7), 3531-60.

Rendall, Michelle. 2010. "Brain Versus Brawn: The Realization of Women's Comparative Advantage." *Mimeo, University of Zurich.*

Rosenzweig, Mark and Junsen Zhang. 2013. "Economic Growth, Comparative Advantage, and Gender Differences in Schooling Outcomes: Evidence from the Birthweight Differences of Chinese Twins." *Journal of Development Economics*, 104, 245-60.

Round, J. M., D. A. Jones, J. W. Honour and A. M. Nevill. 1999. "Hormonal Factors in the Development of Differences in Strength between Boys and Girls during Adolescence: A Longitudinal Study," *Annals of Human Biology* 26(1): 49-62.

Rubalcava, Luis and Graciela Teruel. 2007. "The Mexican Family Life Survey: First Wave," http://www.mxfls.uia.mx,

Ruggles, Steven; J Trent Alexander; Katie Genadek; Ronald Goeken; Matthew B Schroeder and Matthew Sobek. 2010. "Intergrated Public Use Microdata Series: Version 5.0 [Machine-Readable Database]," Minneapolis: University of Minnesota,

Schultz, T Paul. 2004. "School Subsidies for the Poor: Evaluating the Mexican Progresa Poverty Program." *Journal of Development Economics*, 74(1), 199-250.

Sepulveda, J; F Bustreo; R Tapia; J Rivera; R Lozano; G Olaiz; V Partida; MD Garcia and JL Voldespino. 2007. "Improvement of Child Survival in Mexico: The Diagonal Approach." *Salud Publica De Mexico*, 49(S110-S125).

Sepulveda, Jaime; Jose Valdespino and Lourdes Garcia-Garcia. 2006. "Cholera in Mexico: The Paradoxial Benefits of the Last Pandemic." *International Journal of Infectious Diseases*, 10(1), 4-13.

Stein, Aryeh D; Jere R Behrman; Ann DiGirolamo; Ruben Grajeda; Reynaldo Martorell; Agnes Quisumbing and Usha Ramakrishnan. 2005. "Schooling, Educational Achievement, and Cognitive Functioning among Young Guatemalan Adults." *Food and Nutrition Bulletin*, 26(2), S46-S54.

Thomas, Duncan. 1997. "Health and Wages: Evidence on Men and Women in Urban Brazil." *Journal of Econometrics*, 77(1), 159-85.

Tome, Patricia; Hortensia Reyes; Cesar Pina; Leticia Rodriguez and Gonzalo Gutierrez. 1997. "Caracteristicas Asociadas Al Subregistro De Muerto En Ninos Del Estado De Guerrero, Mexico." *Salud Publica Mex*, 39, 523-29.

UNESCO (2010): "Global Education Digest 2010: Comparing Education Statistics Across the World," Discussion paper, UNESCO Institute for Statistics

Venkataramani, Atheendar. 2012. "Early Life Exposure to Malaria and Cognition in Adulthood: Evidence from Mexico." *Journal of Health Economics*, 31, 767-80.

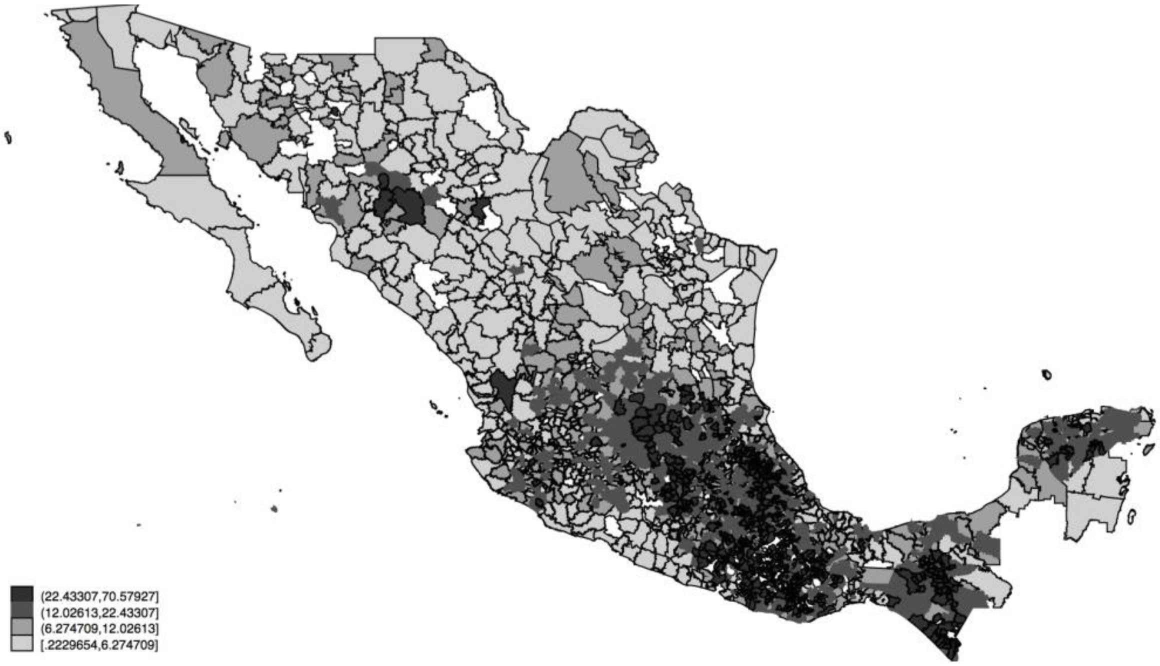
Venkataramani, Atheendar 2009. "The Long-Run and Intergenerational Effects of Early Life Experiences: Evidence from Developing Countries," *Department of Health Policy and Management.* New Haven, CT: Yale University,

Venkataramani, Atheendar and Sonia Bhalotra. 2013. "Effect of Early Life Exposure to a Clean Water Intervention on Adult Stature and Educational Attainment in Mexico: A Quasi-Experimental Retrospective Cohort Study." *Mimeo, University of Essex.*

Vogl, Thomas. 2012. "Height, Skills, and Labor Market Outcomes in Mexico." *NBER Working Paper w18318.*

Watson, Tara. 2006. "Public Health Investments and the Infant Mortality Gap: Evidence from Federal Sanitation Interventions on Us Indian Reservations." *Journal of Public Economics*, 90(8-9), 1537-60.

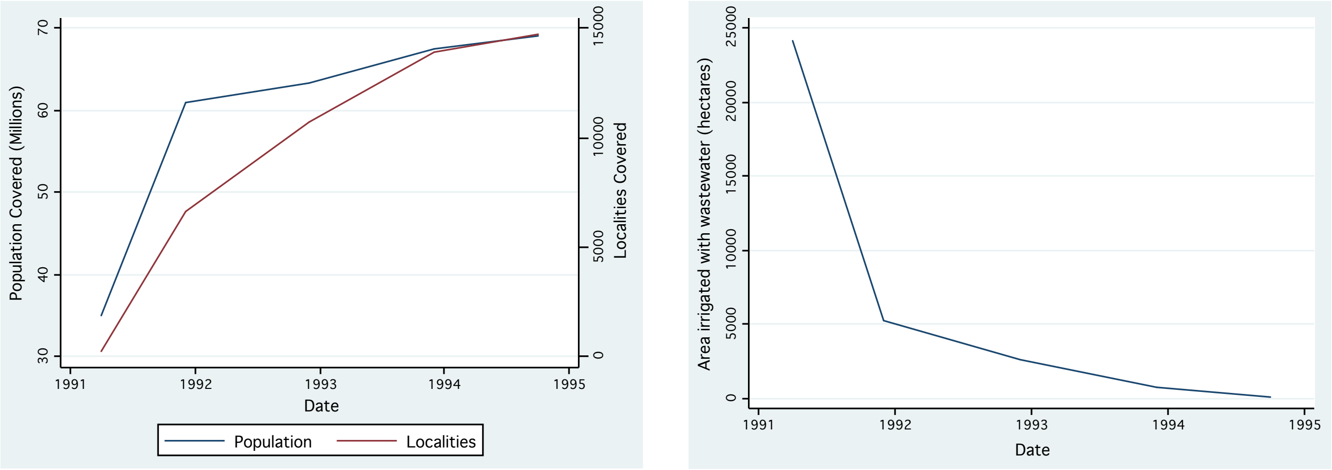
Figure 1 – Pre-Intervention Under-5 Diarrheal Mortality Rates Across Mexican Municipalities (1985-1990 average)



Notes: Map plots average diarrheal mortality rates per 1,000 live births over the period 1985-1990 for children under the age of 5 by municipality. Darker colors reflect higher average pre-intervention diarrheal mortality rates. Data to construct map were obtained from the Mexico Ministry of Health, Vital Statistics.

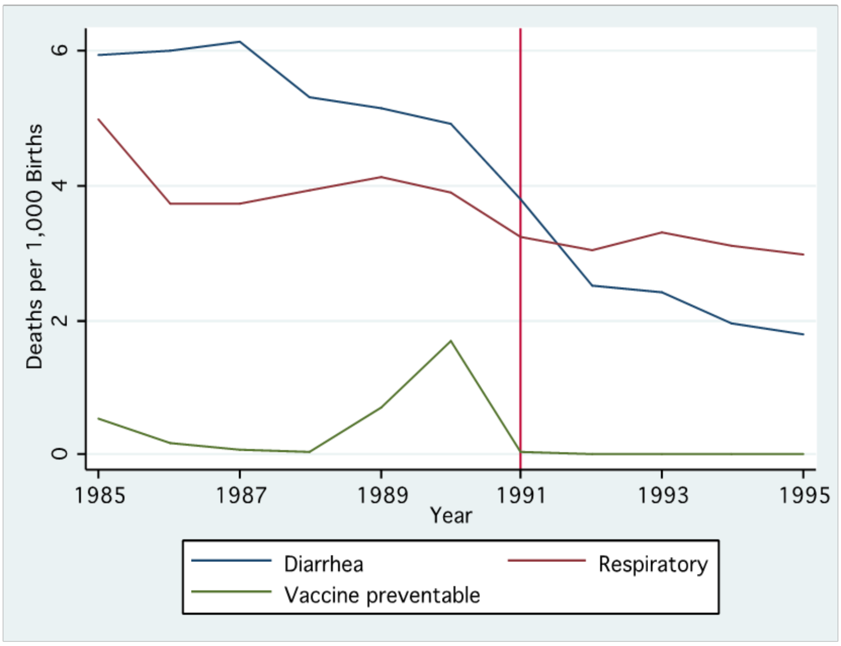
Figure 2 – Timeline and Scope of Mexico's National Clean Water Program

A. Population Access to Chlorinated Water B. Land Area Irrigated with Waste Water



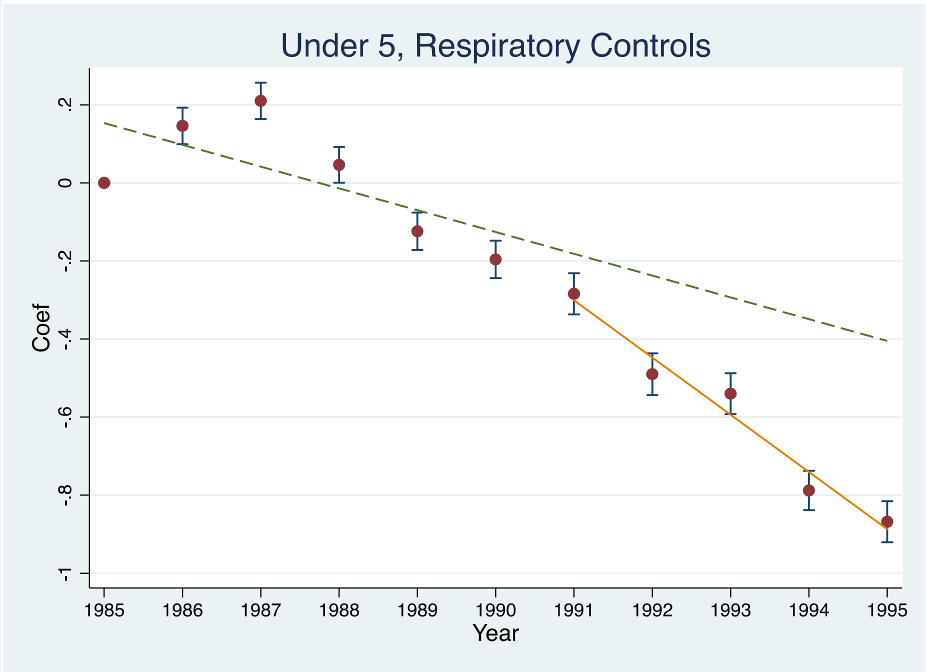
Source: Government of Mexico, National Water Commission

Figure 3 – Trends in Child Infectious Disease Mortality: Diarrhea, Respiratory Infections, and Vaccine Preventable Infections



Source: See *Figure 1*.

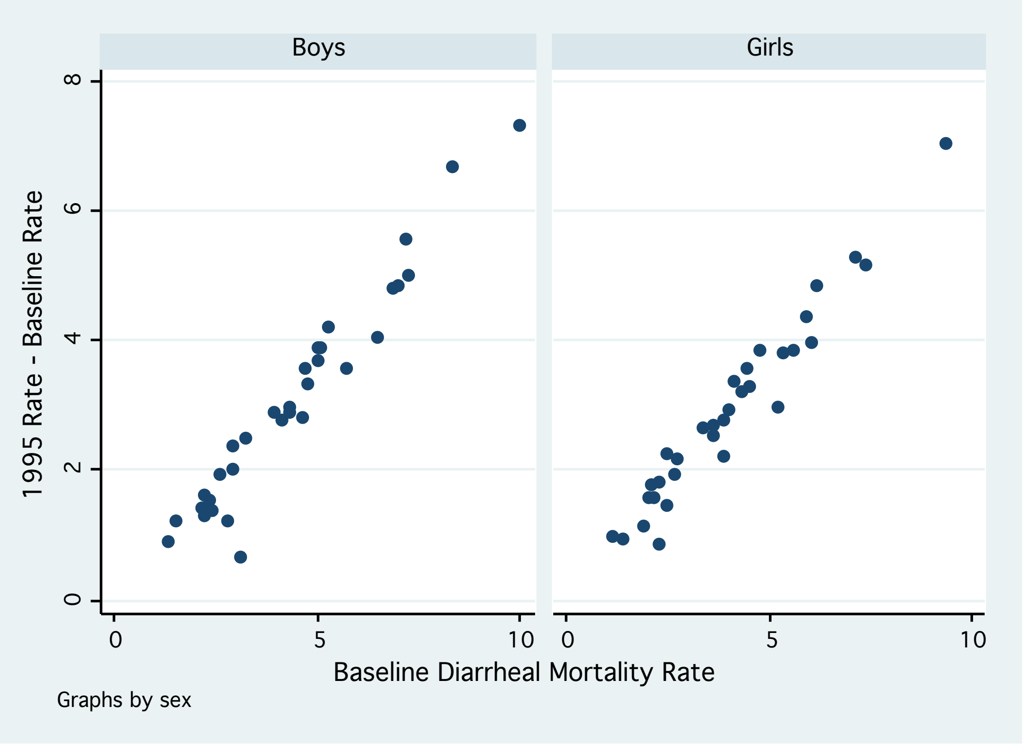
Figure 4 – Event Study Coefficient Plots



Notes: Estimates of the Diarrhea*Year coefficients from Equation 1. Coefficients are denoted by dots and the bands denote the 95% confidence interval of the estimates. The dashed green line is a predicted trend line calculated using pre-1991 coefficients, which we extend across the time series. The solid orange line is the trend line for the 1991-1995 coefficients. The gap between the two can be interpreted as the PAL treatment effect.

Figure 5 – Convergence across States in Child Diarrhea Mortality Rates

Scatter Plot of Absolute Change in Diarrheal Mortality Post-1991 and the Pre-Intervention Rate



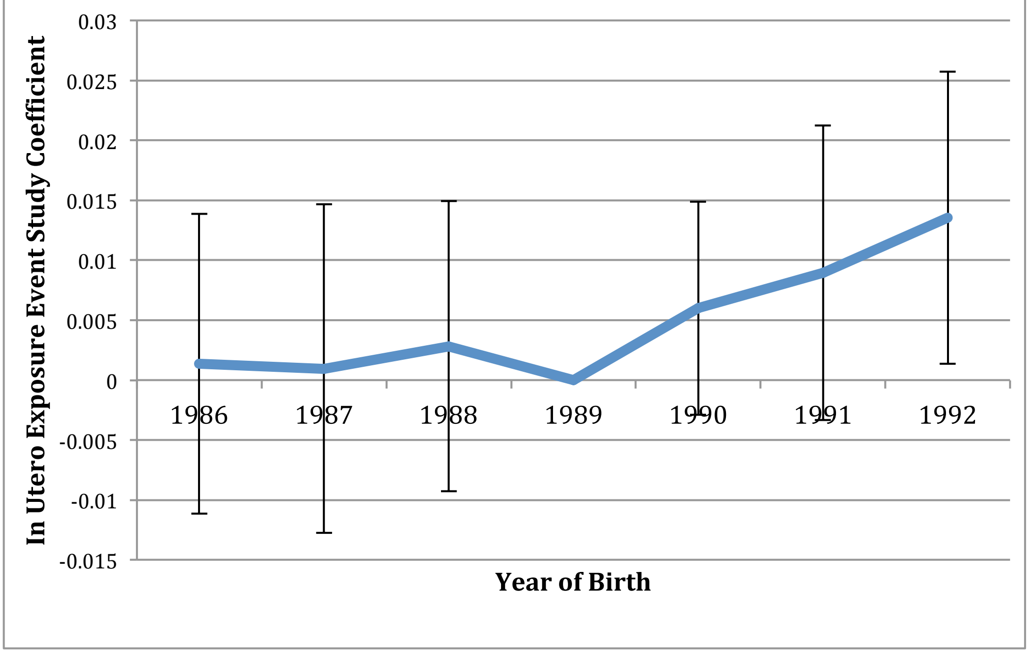


Figure 6: Event Study for Height-for-Age Z-Score

Notes: Standard errors are clustered at the municipality of residence level. 90% confidence intervals are provided. Birth cohorts are defined from 4/1/YOB-3/31/YOB+1.

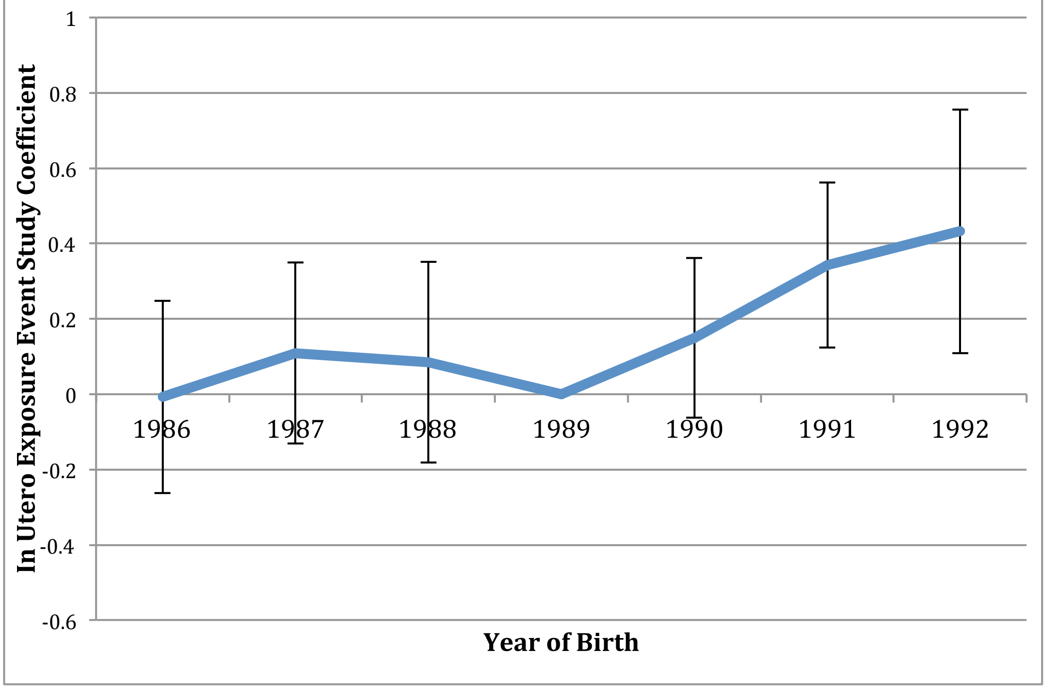


Figure 7: Event Study for Raven's Score

Notes: Standard errors are clustered at the municipality of residence level. 90% confidence intervals are provided. Birth cohorts are defined from 4/1/YOB-3/31/YOB+1.

Table 1: Diarrheal Mortality Relative to Respiratory Diseases Mortality Pre-Post PAL From 1985 and 1995

	(1)	(2)	(3)	(4)
	Under-5 Years	*0-1 Months*	*1-12 Months*	*1-4 Years*
1(Diarrhea)*1(Post)	-0.119***	-0.0651***	-0.136***	0.008
	(0.020)	(0.020)	(0.021)	(0.020)
1(Diarrhea)*1(Post)*Year	-0.0909***	-0.0868***	-0.0422***	-0.0321***
	(0.008)	(0.008)	(0.008)	(0.008)
1(Diarrhea)*Year	-0.0558***	-0.0489***	-0.0509***	-0.0347***
	(0.005)	(0.005)	(0.005)	(0.004)
1(Diarrhea)	-0.133***	-0.361***	0.147***	0.144***
	(0.039)	(0.040)	(0.031)	(0.025)
Observations	48,906	48,906	48,906	48,906
% Decline by 1995 Due to PAL	48.2	41.2	30.5	17.1

Notes: Estimates of Equation 2 in the main text. *** - p<0.01, ** - p<0.05, * - p<0.1. Standard errors are clustered at the municipality level. All models include municipality and year fixed effects and are weighted by municipality baseline live births (i.e., the average for 1988-1990). Each column represents a separate regression, with the column header denoting the specific age group over which the inverse hyperbolic sine transformation of mortality was calculated for the dependent variable. 1(Diarrhea) =1 denotes diarrheal mortality rates, while 1(Diarrhea) = 0 denotes mortality from respiratory diseases. Post = 1 if the year of observation is 1991 or thereafter. The final row, % Decline by 1995 Due to PAL is calculated by adding the coefficient on the level break to four times that of the trend break (since 1995 is 4 years after PAL). The year variable is rescaled such that 1991 = 0 and increments above and below are denoted by positive and negative integers.

Table 2: Impact of Clean Water Program Exposure in Infancy on Height and Raven's Cognitive Score Measured in MxFLS1 For Individuals Born Between 1986 and 1992

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Height for Age Z-Score						
Full*Base Rate	-0.016***	0.009***	0.011*	0.011*	0.017***	0.015**
	[0.003]	[0.003]	[0.006]	[0.006]	[0.006]	[0.006]
Partial*Base Rate	-0.022***	0.001	0.002	0.006	0.009	0.006
	[0.004]	[0.004]	[0.005]	[0.005]	[0.006]	[0.007]
Mean dep. variable	-0.43	-0.43	-0.43	-0.56	-0.56	-0.56
Observations	4,687	4,687	4,687	2,726	2,726	2,724
Number of mothers	-	-	-	1,172	1,172	1,171
Panel B: Raven's % Score (0-100)						
Full*Base Rate	-0.203***	0.091**	0.171*	0.340***	0.313***	0.343***
	[0.059]	[0.044]	[0.103]	[0.113]	[0.104]	[0.120]
Partial*Base Rate	-0.278***	0.036	0.063	0.082	0.211*	0.324***
	[0.065]	[0.043]	[0.078]	[0.112]	[0.117]	[0.121]
Mean dep. variable	61.74	61.74	61.74	60.40	60.40	60.39
Observations	4,973	4,973	4,973	2,909	2,909	2,907
Number of mothers	-	-	-	1,254	1,254	1,253
Municipality Fixed Effects	NO	YES	YES	YES	YES	YES
Municipality Specific Trends	NO	NO	YES	YES	YES	YES
Maternal Fixed Effects	NO	NO	NO	YES	YES	YES
Post*Pre-Intervention Non-Diarrheal Child Mortality Rates	NO	NO	NO	NO	YES	YES
State of Residence-by-Year of Birth Fixed Effects	NO	NO	NO	NO	NO	YES

Notes: Standard errors are clustered at the municipality of residence level. *** p<0.01, ** p<0.05, * p<0.1. All specifications include an indicator for gender, date of birth fixed effects, date of interview fixed effects, birth order fixed effects, mother's age at birth, and mother's age at birth squared. Raven Progressive Matrix test scores are defined as the fraction of questions answered correctly, range 0 to 100. Height for age Z-Score calculated using the STATA command "zanthro" and utilizing the WHO child growth charts. "Full"=1 if the individual was born on or after April 1991, "Partial"=1 if the individual was born between April 1990 and March 1991, and "Base Rate" is the gender-specific child diarrheal mortality rate in the municipality of residence averaged across the pre-intervention years, 1988-1990. "Municipality Specific Trends" includes municipality-specific linear time trends and two vectors of pre-intervention municipality of residence characteristics interacted with "Full" and "Partial", respectively. Pre-intervention municipality characteristics include: percent of population with access to piped water, percent of population with access to sewage system, percent of population has indigenous heritage, mean education level, and mean log of income. Pre-intervention non-diarrheal child mortality rates include: respiratory diseases and vaccine preventable diseases.

Table 3: Impact of Clean Water Program Exposure in Infancy on Height and Raven's Cognitive Score Measured in MxFLS1
For Individuals Born Between 1986 and 1992
Assigning Location Using Available Municipality of Birth Information

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Height for Age Z-Score						
Full*Base Rate	-0.016***	0.008**	0.012*	0.009	0.016*	0.016*
	[0.003]	[0.003]	[0.007]	[0.010]	[0.009]	[0.009]
Partial*Base Rate	-0.020***	0.002	0.005	0.009	0.009	0.008
	[0.004]	[0.004]	[0.005]	[0.007]	[0.009]	[0.011]
Mean dep. variable	-0.45	-0.45	-0.45	-0.58	-0.58	-0.58
Observations	4,226	4,136	4,136	2,344	2,344	2,335
Number of mothers	-	-	-	1,023	1,023	1,019
Panel B: Raven's % Score (0-100)						
Full*Base Rate	-0.206***	0.065	0.174	0.360**	0.342**	0.406**
	[0.061]	[0.049]	[0.107]	[0.147]	[0.141]	[0.173]
Partial*Base Rate	-0.265***	0.055	0.122	0.105	0.213	0.319
	[0.065]	[0.041]	[0.078]	[0.166]	[0.178]	[0.198]
Mean dep. variable	61.48	61.50	61.50	60.09	60.09	60.08
Observations	4,492	4,397	4,397	2,495	2,495	2,486
Number of mothers	-	-	-	1,088	1,088	1,084
Municipality of Birth Fixed Effects	NO	YES	YES	YES	YES	YES
Municipality Specific Trends	NO	NO	YES	YES	YES	YES
Maternal Fixed Effects	NO	NO	NO	YES	YES	YES
Post*Pre-Intervention Non-Diarrheal Child Mortality Rates	NO	NO	NO	NO	YES	YES
State of Birth-by-Year of Birth Fixed Effects	NO	NO	NO	NO	NO	YES

Notes: Standard errors are clustered at the municipality of birth level. *** p<0.01, ** p<0.05, * p<0.1. All specifications include an indicator for gender, date of birth fixed effects, date of interview fixed effects, birth order fixed effects, mother's age at birth, and mother's age at birth squared. Raven Progressive Matrix test scores are defined as the fraction of questions answered correctly, range 0 to 100. Height for age Z-Score calculated using the STATA command "zanthro" and utilizing the WHO child growth charts. "Full"=1 if the individual was born on or after April 1991, "Partial"=1 if the individual was born between April 1990 and March 1991, and "Base Rate" is the gender-specific child diarrheal mortality rate in the municipality of birth averaged across the pre-intervention years, 1988-1990. "Municipality Specific Trends" includes municipality-specific linear time trends and two vectors of pre-intervention municipality of birth characteristics interacted with "Full" and "Partial", respectively. Pre-intervention municipality characteristics include: percent of population with access to piped water, percent of population with access to sewage system, percent of population has indigenous heritage, mean education level, and mean log of income. Pre-intervention non-diarrheal child mortality rates include: respiratory diseases and vaccine preventable diseases.

Table 4: Impact of Clean Water Program Exposure in Infancy on Height and Raven's Cognitive Score Measured in MxFLS3
For Individuals Born Between 1986 and 1992

	MxFLS1 Survey		MxFLS3 Survey			
	Full Sample	MxFLS3 Sample				
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Height for Age Z-Score						
Full*Base Rate	0.011*	0.006	0.004	0.011*	0.018***	0.012*
	[0.006]	[0.007]	[0.004]	[0.006]	[0.007]	[0.007]
Partial*Base Rate	0.002	0.004	0.003	0.003	0.002	-0.001
	[0.005]	[0.006]	[0.003]	[0.005]	[0.005]	[0.006]
Mean dep. variable	-0.433	-0.41	0.00	-0.03	-0.03	-0.03
Observations	4,687	3,057	3,057	1,580	1,580	1,576
Number of mothers	-	-	-	708	708	706
Panel B: Raven's % Score (0-100)						
Full*Base Rate	0.171*	0.089	0.104	0.536*	0.452	0.247
	[0.103]	[0.124]	[0.106]	[0.309]	[0.320]	[0.504]
Partial*Base Rate	0.063	-0.017	-0.181	0.136	0.248	0.517
	[0.078]	[0.080]	[0.120]	[0.215]	[0.264]	[0.358]
Mean dep. variable	61.74	63.38	57.50	55.54	55.54	55.46
Observations	4,973	2,726	2,726	1,224	1,224	1,208
Number of mothers	-	-	-	558	558	550
Municipality Fixed Effects	YES	YES	YES	YES	YES	YES
Municipality Specific Trends	YES	YES	YES	YES	YES	YES
Maternal Fixed Effects	NO	NO	NO	YES	YES	YES
Post*Pre-Intervention Non-Diarrheal Child Mortality Rates	NO	NO	NO	NO	YES	YES
State of Birth-by-Year of Birth Fixed Effects	NO	NO	NO	NO	NO	YES

Notes: Standard errors are clustered at the municipality of residence level. *** p<0.01, ** p<0.05, * p<0.1. All specifications include an indicator for gender, date of birth fixed effects, date of interview fixed effects, birth order fixed effects, mother's age at birth, and mother's age at birth squared. Raven Progressive Matrix test scores are defined as the fraction of questions answered correctly, range 0 to 100. Height-for-age z-Score in MxFLS1 calculated using the STATA command "zanthro" and utilizing the WHO child growth charts. Height for age z-scores in MxFLS3 are calculated as (height-mean height)/standard deviation of height separately for each age. "Full"=1 if the individual was born on or after April 1991, "Partial"=1 if the individual was born between April 1990 and March 1991, and "Base Rate" is the gender-specific child diarrheal mortality rate in the municipality of residence averaged across the pre-intervention years, 1988-1990. "Municipality Specific Trends" includes municipality-specific linear time trends and two vectors of pre-intervention municipality of residence characteristics interacted with "Full" and "Partial", respectively. Pre-intervention municipality characteristics include: percent of population with access to piped water, percent of population with access to sewage system, percent of population has indigenous heritage, mean education level, and mean log of income. Pre-intervention non-diarrheal child mortality rates include: respiratory diseases and vaccine preventable diseases.

Table 5: Impact of Clean Water Program Exposure in Infancy on Other Human Capital Outcomes Measured in MxFLS1 For Individuals Born Between 1986 and 1992

	Years of Education (1)	Attends School (2)	Repeated Grade (3)	Started Secondary School (4)	Employed (5)
Full*Base Rate	0.020**	-0.001	-0.008*	0.005**	-0.005***
	[0.009]	[0.002]	[0.004]	[0.002]	[0.002]
Partial*Base Rate	0.007	0.000	-0.002	0.003	-0.004**
	[0.009]	[0.002]	[0.003]	[0.002]	[0.002]
Mean dep. variable	5.31	0.895	0.274	0.316	0.125
Observations	3,033	2,971	2,926	3,033	3,015
Number of mothers	1,292	1,278	1,262	1,298	1,292
Municipality Fixed Effects	YES	YES	YES	YES	YES
Municipality Specific Trends	YES	YES	YES	YES	YES
Maternal Fixed Effects	YES	YES	YES	YES	YES

Notes: Standard errors are clustered at the municipality of residence level. *** p<0.01, ** p<0.05, * p<0.1. All specifications include date of birth fixed effects, date of interview fixed effects, birth order fixed effects, mother's age at birth, and mother's age at birth squared. "Full"=1 if the individual was born on or after April 1991, "Partial"=1 if the individual was born between April 1990 and March 1991, and "Base Rate" is the gender-specific child diarrheal mortality rate in the municipality of residence averaged across the pre-intervention years, 1988-1990. "Municipality Specific Trends" includes municipality-specific linear time trends and two vectors of pre-intervention municipality of residence characteristics interacted with "Full" and "Partial", respectively. Pre-intervention municipality characteristics include: percent of population with access to piped water, percent of population with access to sewage system, percent of population has indigenous heritage, mean education level, and mean log of income.

Table 6: Impact of Clean Water Program Exposure in Infancy on Economic Outcomes Measured in MxFLS3 For Individuals Born Between 1986 and 1992

	Years of Education (1)	Employed (2)	Hours Worked Last Week Conditional on Working (3)	Earnings per Hour (4)
Full*Base Rate	0.011	0.000	-0.503	2.389*
	[0.014]	[0.005]	[0.822]	[1.221]
Partial*Base Rate	0.017	0.002	-0.455	-0.657
	[0.014]	[0.003]	[0.621]	[1.526]
Mean dep. variable	9.70	0.47	42.46	23.12
Observations	2,900	2,423	640	556
Number of mothers	1,244	1,072	302	263
Municipality Fixed Effects	YES	YES	YES	YES
Municipality Specific Trends	YES	YES	YES	YES
Maternal Fixed Effects	YES	YES	YES	YES

Notes: Standard errors are clustered at the municipality of residence level. *** p<0.01, ** p<0.05, * p<0.1. All specifications include date of birth fixed effects, date of interview fixed effects, birth order fixed effects, mother's age at birth, and mother's age at birth squared. "Full"=1 if the individual was born on or after April 1991, "Partial"=1 if the individual was born between April 1990 and March 1991, and "Base Rate" is the gender-specific child diarrheal mortality rate in the municipality of residence averaged across the pre-intervention years, 1988-1990. "Municipality Specific Trends" includes municipality-specific linear time trends and two vectors of pre-intervention municipality of residence characteristics interacted with "Full" and "Partial", respectively. Pre-intervention municipality characteristics include: percent of population with access to piped water, percent of population with access to sewage system, percent of population has indigenous heritage, mean education level, and mean log of income.

Table 7: Impact of Clean Water Program Exposure in Infancy on Height and Raven's Cognitive Score Measured in MxFLS1 by Gender For Individuals Born Between 1986 and 1992

	(1)	(2)	(3)	(4)	(5)
Panel A: Height for Age Z-Score					
Full*Base Rate	0.023***	0.026***	0.024***	0.028***	0.027***
	[0.004]	[0.006]	[0.007]	[0.008]	[0.009]
Full*Base Rate*I(Male=1)	-0.021***	-0.012**	-0.009	-0.004	-0.001
	[0.005]	[0.006]	[0.008]	[0.008]	[0.009]
p-value for H0: Total "Full" Male Effect=0	0.72	0.09	0.01	0.00	0.00
Mean dep. variable	-0.43	-0.43	-0.56	-0.56	-0.56
Observations	4,687	4,687	2,726	2,726	2,724
Number of mothers	-	-	1,172	1,172	1,171
Panel B: Raven's % Score (0-100)					
Full*Base Rate	0.101	0.306**	0.522***	0.451***	0.527***
	[0.070]	[0.135]	[0.134]	[0.140]	[0.158]
Full*Base Rate*I(Male=1)	-0.017	-0.214*	-0.376**	-0.325**	-0.324**
	[0.070]	[0.123]	[0.163]	[0.152]	[0.157]
p-value for H0: Total "Full" Male Effect=0	0.07	0.60	0.38	0.41	0.23
Mean dep. variable	61.74	61.74	60.40	60.40	60.39
Observations	4,973	4,973	2,909	2,909	2,907
Number of mothers	-	-	1,254	1,254	1,253
Municipality Fixed Effects	YES	YES	YES	YES	YES
Municipality Specific Trends	NO	YES	YES	YES	YES
Maternal Fixed Effects	NO	NO	YES	YES	YES
Post*Pre-Intervention Non-Diarrheal Child Mortality Rates	NO	NO	NO	YES	YES
State of Birth-by-Year of Birth Fixed Effects	NO	NO	NO	NO	YES

Notes: Standard errors are clustered at the municipality of residence level. *** p<0.01, ** p<0.05, * p<0.1. All specifications include an indicator for gender, date of birth fixed effects, date of interview fixed effects, birth order fixed effects, mother's age at birth, and mother's age at birth squared. Raven Progressive Matrix test scores are defined as the fraction of questions answered correctly, range 0 to 100. Height for age Z-Score calculated using the STATA command "zanthro" and utilizing the WHO child growth charts. "Full"=1 if the individual was born on or after April 1991, "Partial"=1 if the individual was born between April 1990 and March 1991, and "Base Rate" is the gender-specific child diarrheal mortality rate in the municipality of residence averaged across the pre-intervention years, 1988-1990. "Municipality Specific Trends" includes municipality-specific linear time trends and two vectors of pre-intervention municipality of residence characteristics interacted with "Full" and "Partial" for each gender, respectively. Pre-intervention municipality characteristics include: percent of population with access to piped water, percent of population with access to sewage system, percent of population has indigenous heritage, mean education level, and mean log of income. Pre-intervention non-diarrheal child mortality rates include: respiratory diseases and vaccine preventable diseases.

Table 8: Impact of Clean Water Program Exposure in Infancy on Other Human Capital Outcomes Measured in MxFLS1 by Gender For Individuals Born Between 1986 and 1992

	Years of Education (1)	Attends School (2)	Repeated Grade (3)	Started Secondary School (4)	Employed (5)
Full*Base Rate	0.007	0.000	-0.006	0.004	-0.004
	[0.012]	[0.002]	[0.005]	[0.003]	[0.003]
Full*Base Rate*I(Male=1)	0.027***	-0.001	-0.004	0.003	-0.002
	[0.010]	[0.001]	[0.004]	[0.002]	[0.002]
p-value for H0: Total "Full" Male Effect=0	0.00	0.60	0.01	0.00	0.00
Mean dep. variable	5.31	0.895	0.274	0.316	0.125
Observations	3,033	2,971	2,926	3,033	3,015
Number of mothers	1,292	1,278	1,262	1,298	1,292
Municipality Fixed Effects	YES	YES	YES	YES	YES
Municipality Specific Trends	YES	YES	YES	YES	YES
Maternal Fixed Effects	YES	YES	YES	YES	YES

Notes: Standard errors are clustered at the municipality of residence level. *** p<0.01, ** p<0.05, * p<0.1. All specifications include an indicator for gender, date of birth fixed effects, date of interview fixed effects, birth order fixed effects, mother's age at birth, and mother's age at birth squared. "Full"=1 if the individual was born on or after April 1991, "Partial"=1 if the individual was born between April 1990 and March 1991, and "Base Rate" is the gender-specific child diarrheal mortality rate in the municipality of residence averaged across the pre-intervention years, 1988-1990. "Municipality Specific Trends" includes municipality-specific linear time trends and two vectors of pre-intervention municipality of residence characteristics interacted with "Full" and "Partial" for each gender, respectively. Pre-intervention municipality characteristics include: percent of population with access to piped water, percent of population with access to sewage system, percent of population has indigenous heritage, mean education level, and mean log of income.

Table 9: Impact of Clean Water Program Exposure in Infancy on Economic Outcomes Measured in MxFLS3 by Gender For Individuals Born Between 1986 and 1992

	Years of Education (1)	**Employed** (2)	**Hours Worked Last Week Conditional on Working** (3)	**Earnings per Hour** (4)
Full*Base Rate	0.020	0.003	-0.590	4.664*
	[0.019]	[0.006]	[1.297]	[2.375]
Full*Base Rate*I(Male=1)	-0.008	0.000	-0.111	0.061
	[0.018]	[0.006]	[1.342]	[3.459]
p-value for H0: Total "Full" Male Effect=0	0.43	0.58	0.48	0.02
Mean dep. variable	9.70	0.47	42.46	23.12
Observations	2,900	2,423	640	556
Number of mothers	1,244	1,072	302	263
Municipality Fixed Effects	YES	YES	YES	YES
Municipality Specific Trends	YES	YES	YES	YES
Maternal Fixed Effects	YES	YES	YES	YES

Notes: Standard errors are clustered at the municipality of residence level. *** p<0.01, ** p<0.05, * p<0.1. All specifications include an indicator for gender, date of birth fixed effects, date of interview fixed effects, birth order fixed effects, mother's age at birth, and mother's age at birth squared. "Full"=1 if the individual was born on or after April 1991, "Partial"=1 if the individual was born between April 1990 and March 1991, and "Base Rate" is the gender-specific child diarrheal mortality rate in the municipality of residence averaged across the pre-intervention years, 1988-1990. "Municipality Specific Trends" includes municipality-specific linear time trends and two vectors of pre-intervention municipality of residence characteristics interacted with "Full" and "Partial" for each gender, respectively. Pre-intervention municipality characteristics include: percent of population with access to piped water, percent of population with access to sewage system, percent of population has indigenous heritage, mean education level, and mean log of income.

Appendix Table 1: Impact of Clean Water Program Exposure in Infancy on Height and Raven's Cognitive Score Measured in MxFLS1 For Individuals Born Between 1986 and 1992 Using Maternal Fixed Effects Sample

	(1)	(2)	(3)
Panel A: Height for Age Z-Score			
Full*Base Rate	-0.013***	0.011**	0.016*
	[0.004]	[0.004]	[0.008]
Partial*Base Rate	-0.020***	0.001	0.010*
	[0.004]	[0.003]	[0.006]
Mean dep. variable	-0.56	-0.56	-0.56
Observations	2,726	2,726	2,726
Panel B: Raven's % Score (0-100)			
Full*Base Rate	-0.154**	0.101	0.108
	[0.067]	[0.067]	[0.105]
Partial*Base Rate	-0.224***	0.044	0.121
	[0.078]	[0.072]	[0.088]
Mean dep. variable	60.40	60.40	60.40
Observations	2,909	2,909	2,909
Municipality Fixed Effects	NO	YES	YES
Municipality Specific Trends	NO	NO	YES

Notes: Standard errors are clustered at the municipality of residence level. *** p<0.01, ** p<0.05, * p<0.1. All specifications include an indicator for gender, date of birth fixed effects, date of interview fixed effects, birth order fixed effects, mother's age at birth, and mother's age at birth squared. Raven Progressive Matrix test scores are defined as the fraction of questions answered correctly, range 0 to 100. Height for age Z-Score calculated using the STATA command "zanthro" and utilizing the WHO child growth charts. "Full"=1 if the individual was born on or after April 1991, "Partial"=1 if the individual was born between April 1990 and March 1991, and "Base Rate" is the gender-specific child diarrheal mortality rate in the municipality of residence averaged across the pre-intervention years, 1988-1990. "Municipality Specific Trends" includes municipality-specific linear time trends and two vectors of pre-intervention municipality of residence characteristics interacted with "Full" and "Partial", respectively. Pre-intervention municipality characteristics include: percent of population with access to piped water, percent of population with access to sewage system, percent of population has indigenous heritage, mean education level, and mean log of income.

Appendix Table 2: Impact of Clean Water Program Exposure in Infancy on Mother's Characteristics Measured in MxFLS1 For Individuals Born Between 1986 and 1992

	Mother's Characteristics:					
	Raven's Score	**Height for Age Z-Score**	**Household Per Capita Expenditure**	**Underweight (BMI<18.5)**	**Obese (BMI>30)**	**Finished Compulsory Schooling**
	(1)	(2)	(3)	(4)	(5)	(6)
Full*Base Rate	-0.148	0.001	-24.317*	0.002*	0.003	0.001
	[0.096]	[0.006]	[13.241]	[0.001]	[0.002]	[0.002]
Partial*Base Rate	-0.074	0.010***	-6.179	0.001***	-0.001	0.000
	[0.097]	[0.003]	[10.999]	[0.001]	[0.002]	[0.001]
Mean dep. variable	43.54	-0.03	1,165.10	0.007	0.374	0.317
Observations	4,579	4,465	5,066	4,425	4,425	4,814
Municipality Fixed Effects	YES	YES	YES	YES	YES	YES
Municipality Specific Trends	YES	YES	YES	YES	YES	YES
Post*Pre-Intervention Non-Diarrheal Child Mortality	YES	YES	YES	YES	YES	YES
State of Residence-by-Year of Birth Fixed Effects	YES	YES	YES	YES	YES	YES

Notes: Standard errors are clustered at the municipality of residence level. *** p<0.01, ** p<0.05, * p<0.1. All specifications include date of birth fixed effects, date of interview fixed effects, birth order fixed effects, mother's age at birth, and mother's age at birth squared. Raven Progressive Matrix test scores are defined as the fraction of questions answered correctly, range 0 to 100. Height for age z-scores are calculated as (height-mean height)/standard deviation of height separately for each age. "Full"=1 if the individual was born on or after April 1991, "Partial"=1 if the individual was born between April 1990 and March 1991, and "Base Rate" is the gender-specific child diarrheal mortality rate in the municipality of residence averaged across the pre-intervention years, 1988-1990. "Municipality Specific Trends" includes municipality-specific linear time trends and two vectors of pre-intervention municipality of residence characteristics interacted with "Full" and "Partial", respectively. Pre-intervention municipality characteristics include: percent of population with access to piped water, percent of population with access to sewage system, percent of population has indigenous heritage, mean education level, and mean log of income. Pre-intervention non-diarrheal child mortality rates include: respiratory diseases and vaccine preventable diseases.