ISSN 1755-5361

University of Essex

Department of Economics

Discussion Paper Series

No. 745 December 2013

Cognitive Development and Infectious Disease: Gender Differences in Investments and Outcomes

Sonia R. Bhalotra and Atheendar S. Venkataramani

Cognitive Development and Infectious Disease:
Gender Differences in Investments and Outcomes

13 December 2013

Sonia R. Bhalotra
Department of Economics
University of Bristol, UK
s.bhalotra@bristol.ac.uk

Atheendar S. Venkataramani
Department of Medicine
Massachusetts General Hospital, USA
avenkataramani@partners.org

Abstract: We exploit exogenous variation in the risk of waterborne disease created by implementation of a major water reform in Mexico in 1991 to investigate impacts of infant exposure on indicators of cognitive development and academic achievement in late childhood. We estimate that a one standard deviation reduction in childhood diarrhea mortality rates results in about a 0.1 standard deviation increase in test scores, but only for girls. We show that a reason for the gender differentiated impacts is that the water reform induces parents to make complementary investments in education that favor girls, consistent with their comparative advantage in skilled occupations. The results provide novel evidence of the potential for clean water provision to narrow test score gaps across countries and, within countries, across gender.

Acknowledgements: This paper has benefited from presentation at the AEA meetings in San Diego and at seminars at the Paris School of Economics, the Institute of Fiscal Studies in London and the Universities of Pompeau Fabra, Bristol, Oxford, Sussex and Essex. We received helpful comments from Orazio Attanasio, Manuel Bagues, Erlend Berg, Antonio Ciccone, James Fenske, Paul Schultz, Marcos vera Hernandez and Simon Quinn. Damian Clark, Peter Gerrish, Claudia Herresthal, Salomo Hirvonen, and Yichau Wu provided excellent research assistance. We would like to thank Tania Barham and Ernesto Aguayo-Tellez for generously making available their data on *Opportunidades* and sectoral shares of Mexican state-level GDP, respectively. All errors are our own.

I. Introduction

A growing literature underlines the importance of cognitive ability and achievement in driving educational attainment and labor market outcomes (Heckman et al 2006) and the emerging consensus is that early childhood represents a critical period for the formation of cognitive skills (Grantham-McGregor et al. 2007; Heckman 2007, Cunha and Heckman 2007). This paper focuses on the role of early childhood health in supporting the rapid pace of brain development, the high degree of neuronal plasticity that mark this stage of life, and its eventual influence on skill formation. In particular, severe or repeated early life infections may divert nutrients away from neurological development, particularly during infancy, when it is estimated that about 85% of calorie intake is used to build brains (Eppig et al 2010).[1] 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). 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 explain their weaker performance on international intelligence tests. Similarly, the secular rise in intelligence scores, the "Flynn effect" (James Flynn 1984; Flynn 1987) is consistent with progression through the epidemiological transition (The Economist 2010; Eppig et al 2010). 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

---

[1] By the age of two, children have typically attained 90% of their adult head size. Brain activity consumes 44% of a 5-year old child's metabolic intake and the figure for adults is estimated at 25% (Eppig et al 2010).

infections and later life cognition is at best suggestive, and we know little about the size of any biological impacts and their dynamic consequences.

This paper attempts to fill that niche. It examines the impact of a policy that dramatically improved access to clean water on indicators of cognitive development among a sample of young Mexicans. 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, 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 the extent that the clean water reform improved health and cognitive endowments in the infant period, this will tend to have raised the marginal productivity of subsequent investments in reform-treated cohorts and, by reducing morbidity and mortality from diarrhea, it will have extended the horizon over which returns flow (Heckman and Cunha 1997, Soares 2005). Any measured medium or longer term cognitive gains will incorporate both the stronger infant endowments and any reinforcing parental investments induced by the water reform. We model these investments and attempt to identify the importance of behavioral relative to biological impacts by exploiting gender variation in labor market returns to investments determined by a biologically premised (and hence exogenously given) gender-specific comparative advantage in brain *vs* brawn-intensive tasks.

To account for selectivity into infection, we exploit the introduction of the *Programa de Agua Limpia* (National Clean Water Program) in Mexico, a large-scale nationwide effort, the introduction of which was plausibly exogenous as it was introduced quite suddenly in reaction to the threat of cholera created by an epidemic spreading through the countries neighboring Mexico. We show that introduction of the program led to rapid and sizeable drops in childhood diarrhea mortality rates. Since the absolute decline was systematically larger in states with higher pre-program diarrhea mortality rates, our identification strategy exploits birth state and birth cohort variation, combining the sharp convergence across the states with the timing of the nationwide reform. Identification is strengthened by introducing respiratory infections as a placebo or control disease, which is likely to be similarly responsive to generic public health initiatives but less responsive than diarrhea to a clean water initiative.

We find sizeable positive effects of exposure to clean water (and thereby reduced diarrhea) in infancy on performance in Raven tests conducted at the ages of 9-14 and in math and reading tests taken at the age of 15. However these impacts are only robustly identified for girls, and this gender difference in outcomes holds in two different data samples for three different test scores. Girls exhibit a roughly 0.1 standard deviation increase in test scores in response to a one standard deviation decrease in diarrhea mortality rates.[2] Further analysis indicates that the largest impacts flow from birth-month exposure although, as we elaborate below, the long run impacts appear to depend critically upon the initial biological mechanism being reinforced by subsequent investments later in childhood.

In a substantive second section of the paper, we explore alternative mechanisms behind the gender difference in cognitive gains from the water reform. We show that the evidence favors the hypothesis that the larger impacts on cognition for girls derive from their having received stronger parental investments in education that were complementary to (and reinforcing of) their improved health endowments. Gender-differentiated parental responses are consistent with a Roy-type model in which boys have a (biologically-premised) comparative advantage in brawn or, conversely, girls a comparative advantage in brain-intensive occupations, which raises the average return to education for girls relative to boys (Pitt et al 2012; Rendall 2010; Rosenzweig and Zhang 2012).

---

[2] Our finding that the impacts of the water reform are restricted to girls increases confidence in the identification strategy. To generate our results, any omitted variables would have to vary systematically not only post-1991 and with larger impacts in states with higher pre-1991 diarrhea mortality rates (and with post-1991 trends in control diseases held constant) but also by gender.

We show that girls treated by the water reform were significantly more likely to attend preschool and school as children, they were more likely to have educational resources such as a desk or computer at home, and they shifted time towards homework and away from chores. Treated boys were more likely to attend school but did not receive a significant increase in the other investments.[3] We show that macroeconomic trends in Mexico in the period relevant to our sample cohorts were consistent with the Roy model. Mexican women are disproportionately selected into brain-intensive occupations, this sorting is increasing in education, and female labor force participation has been increasing, all against a backdrop of skill-biased technical change. In line with these changes in employment and occupation, women's education was rising more rapidly than that of men.[4]

Our study makes substantive contributions in several domains. We provide the first causal evidence that the benefits of water interventions that limit early childhood diarrhea incidence extend beyond lowering morbidity and mortality to enhancing cognitive skill formation and academic achievement, with potential impacts on productivity. Recent work highlights only the immediate health-related benefits from water interventions, translated into increased disability adjusted life years (DALYs) (Ahuja et al 2010; Cutler and Miller 2005; Galiani et al 2005; Gamper-Rabindran et al 2005; Watson 2005), and the allocation of global health resources is guided by these more conservative benefits. A recent study by Lim et al (2012) suggests that sanitation and unsafe water rank 26th and 33rd respectively among all causes of DALYs globally (despite diarrhea being one of the top five causes of total years of life lost). Our estimates suggest that the evidence routinely presented to policy-makers, by failing to record the

---

[3] Consistent with the model predictions, boys exposed to the water reform as infants were taller as adults, which indicates that increased investments in nutrition (brawn) favored boys (Bhalotra and Venkataramani 2013).

[4] Our companion paper, using recently released data that capture the sample cohorts when they are young adults, shows that women treated by the water reform completed more years of education, but there was no similar impact for men (Venkataramani and Bhalotra 2013).

socioeconomic benefits that accrue in the long run to survivors, substantially underestimates the benefits of water interventions. This is important as almost 1 billion people lack access to safe drinking water worldwide and it is estimated that each year there are more than 2 million deaths from waterborne diseases, with diarrhea being the most prevalent, and the second leading cause of child death in the world (Fischer Walker et al 2013). [5]

Second, our findings suggest that responsive parental investments that reinforce improvements in infant health driven by the water-reform are critical in delivering the ultimate gains in cognitive outcomes. In this way we contribute some of the scarce evidence in support of models of skill formation positing dynamic complementarities across inputs (Cunha and Heckman 2007; Cunha et al 2010).

A third contribution is to extend an emerging literature that seeks to understand gender differences in the impact of health and education interventions in terms of gender-based comparative advantage in skills (Pitt et al 2012; Rosenzweig and Zhang 2012) by virtue of modeling an exogenous (policy-determined) change in the infant health endowment.[6] Our investigation of gender-differentiation in parental investments marks a

---

[5] An association between the number of diarrheal episodes in early childhood and poorer cognitive development has been demonstrated in the medical literature (Fisher 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. There is no previous work in economics on diarrhea and cognition. However, Barham (2012) and Venkataramani (2012) provide well-identified evidence linking other infectious diseases to cognitive outcomes. 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 Almond (2006), Bleakley (2007, 2010), Barreca (2010), Baird,et al (2011), Cutler et al (2010), Lucas (2010), and Bhalotra and Venkataramani (2012). 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).

[6] Pitt et al (2012) estimate endowments as residuals from a production function for body mass, acknowledging that this approach is sensitive to the choice of functional form and potential endogeneity of the inputs. Rosenzweig and Zhang (2012) use within-twin variation in birth weight

departure from most previous work in the separate domain of studies of the long run impacts of early life interventions (see the survey in Almond and Currie 2011).

A fourth contribution of this study lies in providing evidence of the extent to which investments respond to rates of return to education. A vast literature measures rates of return but the elasticity of investment is not known, possibly because variation across individuals in rates of return is rarely observed and sharp changes in returns seldom occur (see Abramitzky and Lavy 2012; Jensen 2010 and 2012). In this study, the water reform creates a quasi-experimental increase in the rate of return to education-related investments by virtue of improving the infant (cognitive and health) endowment. We estimate the resulting increase in investments in preschool and school attendance and homework. We further show that as the relative return was larger for girls these investment responses were larger for girls.

Fifth, we contribute to the literature on test score gaps between boys and girls. We estimate that the increase in math scores flowing from a one standard deviation decrease in diarrheal mortality rates in childhood closes 80% of the gender gap in Mexico, resulting in a Mexican gender gap smaller than the OECD average gap (which is 11 in 500 points). This is striking given that recent work suggests that the widespread relative underperformance by girls in mathematics may be driven by gender stereotypes (Guiso et al. 2008, Fryer and Levitt 2010; Bharadwaj et al 2012) that, at least in poorer countries, result in girls receiving fewer resources (Bharadwaj et al 2012). Our results imply that parental investments matter for test scores and that they occur in a direction that opposes gender stereotypes since the marginal increase in educational resources following the water reform is greater for girls than boys. We draw the first link between the literature on gender test score gaps and the recent literature premised on biological differences across the sexes in brawn.

---

to identify endowment differences but this approach may lack external validity and is prone to bias if there are within-twin (or sibling) spillovers (Adhvaryu and Nyshadham 2012).

Finally, our findings are relevant to a recent debate concerning whether height is merely a proxy for (often unobservable) cognitive development (Case and Paxson 2008) or whether there are direct labor market returns to height (Persico et al 2004) because, alongside the reform-driven cognitive gains unique to girls that we document in this paper, were reform-driven gains in height for boys that we document elsewhere (Bhalotra and Venkataramani 2013). Together, these results suggest, first, that investigation of returns to height *vs* cognition on the labor market may produce biased evidence if it is not stratified by gender and, second, they highlight the unexploited potential for a natural experiment such as we find in the Mexican clean water reform to contribute to the current debate by identifying the role of cognition *vs* height on the labor market.

The remainder of the paper is laid out as follows. In the first part of the paper, we establish the impacts of improved access to clean water on diarrhea reduction (upon introduction of the policy) and cognitive test scores (for cohorts exposed as infants to the policy change and tested in late childhood). Section II discusses the National Clean Water Program and establishes short-run (or "first-stage") impacts on diarrheal disease risk. Section III discusses the empirical strategy and Section IV the data. The results are discussed in Section V. The second part of the paper, starting with Section VI, analyzes gender differences in cognitive impacts of the water reform, investigating mechanisms. Section VII concludes.

II. Mexico's National Clean Water Program

IIa. History and Description

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).[7] Since 1978-1997, public sector efforts have been highly effective in reducing infectious disease deaths (Frenk et al 2003), with credit typically assigned to

---

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

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*) which involved disinfection of previously untreated water supplies (primarily through chlorination) and the reduction of the use of wastewater for irrigation. The total outlay for the program over the period 1991-1994 was 1 billion USD.[8] Implementation was rapid: the percentage of the population with access to chlorinated water almost doubled between 1990 and 1992 and the number of hectares irrigated with sewage water for the purposes of vegetable cultivation declined by nearly 90% (Gutierrez et al 1996). The bulk of the increase in access to disinfected water occurred within the time span of a half-year (54% covered in April 1991 to 85% covered at the end of that year), see *Figure 2a*. Similarly, farmland area irrigated with wastewater declined markedly between April and December 1991, see *Figure 2b*.[9]

IIb. Establishing Program Impacts on Diarrheal Disease Mortality

In this section we formalize the evidence using age, gender and disease-specific mortality rates by state and year to investigate whether there was a trend break and convergence in disease levels after 1991.

Table 1 shows strong evidence of both a level break (coefficient on *Post*) and trend break (coefficient on *Post*Year)*: child diarrhea mortality drops by 0.38 log points which is

---

[8] Personal communication with Dr. Jaime Sepulveda.

[9] 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).

about 23% of the pre-program mean and the pre-program rate of decline of 6.4% per annum switches to a decline of 16.3% per annum post-1991. Importantly, given our focus on gender differences, we find no significant difference between girls and boys in the level or trend breaks and both coefficients are slightly larger for boys. Thus differential program effects by gender are unlikely to explain our finding that the program had larger impacts on test scores for girls than boys.

We re-estimated the trend break model using a difference in difference specification in which respiratory disease is included as a control disease. Respiratory infection was chosen because it was not directly treated by the 1991 reform (it is not a waterborne infection) but rivaled diarrhea as a leading cause of child morbidity and mortality in Mexico during the period of interest and, *aside from water quality*, shares many of the same risk factors (Bhutta et al 2013). The estimated equation is:

$$(1)\quad ln(M_{dst}) = \beta_0 + \beta_1{*}Treated_d{*}Post_t + \beta_2{*}Treated_d{*}Post_t{*}Year_t + \beta_3{*}Treated_d{*}Year_t + \beta_4{*}Post_t{*}Year_t + \beta_5{*}Post_t + \beta_6{*}Treated_d + \lambda_s + e_{dst}$$

where $ln(M)$ is the logarithm of the under-5 mortality rate from disease d in state s and year t, $Treated = 1$ if the disease is diarrhea (treated by the water reform) and 0 if it is a respiratory condition. This specification differences out changes in public health policy and medical technology that may have influenced diseases other than water-borne diseases, and accounts for pre-existing trends in diarrhea disease mortality (through the $Treated{*}Year$ term). It is a conservative specification given that, despite not being treated directly by the water quality reform, respiratory diseases may still track diarrheal diseases, perhaps because the latter is a possible risk factor for the former (Cutler and Miller 2005; Ashraf et al 2013). This more demanding specification also shows a trend break in diarrhea mortality after the implementation of the National Clean Water Program and that this was significantly larger than the post-1991 decline in respiratory disease mortality (β_2 is negative and significant). The magnitude of the break in trend for diarrhea is similar to that in Table 1. *Figure 3* illustrates these results, 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.[10]

Figure 4a shows that the decline in diarrheal disease mortality occasioned by the Clean Water Program was larger in states with higher pre-intervention rates, or that the reform stimulated convergence in diarrhea across the Mexican states. This holds for boys and girls. *Figure 4b* is an alternative representation of convergence, using a binary measure of the pre-program 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).[12] However, there are no state-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).

III. Research Strategy

IIIa. Basic Model

To assess the impacts of reduced infant diarrhea mortality on cognition and academic achievement in late childhood and adolescence, we estimate:

$$(2) \qquad Y_{ijt} = a_0 + a_1 {*} Post_t {*} BaseRate_j + \boldsymbol{\alpha} {*} \boldsymbol{X}_{jt} + \mu_j + \delta_t + e_{ijt}$$

---

[10] We also estimated the first stage equation with flexible coefficients and this specification confirmed a jump in trend in 1991.

[11] We avoid the use of state-level measures of implementation of the program as these are potentially endogenous (and unavailable).

[12] Specifically, in the late 1980s, the number of episodes of diarrhea under the age of 5 was 5.0 per annum. In 1990 and 1991 it was 4.6 and 4.5, respectively. The rates fell to 3.3 and then 2.2 in 1992 and 1993, respectively.

where Y_{ijt} is the test score for individual *i* born in state *j* in year *t*, *Post* = 1 if the individual was an infant aged 0-12 months in April 1991 (i.e., born after April 1990), *BaseRate* is the pre-intervention gender-specific child mortality rate from diarrheal disease in the individual's state of birth, X_{ijt} is a vector of individual and household-level characteristics[13] and μ_j and δ_t are state and year of birth fixed effects that subsume the main effects *Post* and *BaseRate*. The equations are estimated by gender.

The parameter capturing the impact of the intervention is a_1. By interacting the timing of the intervention with the pre-program state-specific diarrhea mortality rate, we exploit state*cohort variation in program intensity, as in Acemoglu and Johnson (2007), Bleakley (2007), Bhalotra and Venkataramani (2012). Equation (2) may be thought of as the reduced form of a system in which test score outcomes are allowed to depend upon infant exposure to diarrheal disease, with the latter instrumented with the sharp arrival of the National Clean Water program, the impact of which is allowed to vary by the pre-intervention burden of diarrheal disease in the birth state. The identifying assumption is that the potential (test score) outcomes of children are uncorrelated with the timing of the water reform, the validity of which we discuss in the next subsection. The parameter a_1 captures the full impact of clean-water induced disease reductions, which includes both biological impacts on mental and physical growth in infancy as well as any subsequent investments made in response to these endowment changes. We do not observe which individuals are at risk of diarrhea or which benefit from the water reform but estimate intent to treat (ITT) effects. The estimated test score gains for individuals who were in fact "treated" will be larger in proportion to the "participation rate".

---

[13] In the Raven score and time investment models the controls are household head's age, gender and education, housing characteristics (type of wall, floor), log consumption expenditure, size of community and locality of residence. In the PISA math, reading and investment models, the controls are mother's and father's schooling attainment and occupational qualification, school size, indicators of available educational materials (school quality), whether the school attended is public or private, and whether the school is in an urban or rural area.

Our definition of the indicator *Post* assumes that only children exposed to the water reform in infancy are treated. This follows from our reasoning, presented in the introduction, which is that the share of nutritional intake directed towards brain development is highest in the first twelve months of life (Eppig et al 2010). However, later on we test this restriction by estimating more flexible specifications which allow for impacts at other ages.

The standard errors are clustered at the birth state level to allow for serial correlation in the outcomes across years within states (Bertrand et al 2004). One of the two survey datasets we use (described below) only samples respondents from 16 of Mexico's 32 states. To address potential over-rejection of the null when the number of clusters is small, we estimate p-values using the wild cluster-T bootstrap method of Cameron et al (2008).

IIIb. Threats to Inference

Estimates of the parameter of interest will be biased if there are state and time-varying omitted variables correlated with both the test score outcomes Y_{ijt} and the program indicator $Post_t * BaseRate_j$. In this section we discuss how we attempt to limit this bias as follows. Natural potential confounders are similarly timed public health and education efforts but, aside from a measles vaccination effort, there were no other public health interventions within a two-year band around the National Clean Water Program (Sepulveda et al 2007). However, prior interventions could play a role in generating biased estimates of water program impacts if they generate differential pre-trends across the states.[14] We assess this, categorizing states by high and low treatability (indexed by pre-program diarrhea, *BaseRate*) and using a sample restricted to pre-program cohorts (1985-1990), regressing test scores on a flexible trend (cohort dummies) interacted with *BaseRate,*

---

[14] 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.

and we find no evidence of convergence in test scores for boys or girls prior to the reform (*Appendix Table 2*, discussed below).

We nevertheless control for relevant state and year varying variables including the logarithm of state per capita GDP, the child mortality rate from two control or placebo diseases, respiratory disease and vaccine-preventable diseases[15], and rainfall. These variables capture state-specific trends in living standards and the health environment. We also include state literacy and grade attainment to capture any variation in school supply variation that may have directly influenced test scores. In order to bias our estimates any changes in income, the disease environment or school investments would have to exhibit a trend break in 1991 (and be larger in states with higher pre-intervention diarrhea mortality rates). Each of the controls is therefore entered with its pre-1991 level interacted with *Post*, so as to more severely test the attribution of the 1991 break to the water reform and diarrhea reduction.[16]

To allow for any other omitted trends, we assess robustness of the results to inclusion of birth region*birth year fixed effects and birth state specific quadratic time trends.[17] Our restriction of the estimation sample to the 1988-1993 birth cohorts, a relatively narrow window, limits the role of omitted trends and excludes the financial

---

[15] 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.

[16] We investigated a specification in which these state*cohort variables appeared unrestricted and there was no significant change in the coefficient of interest.

[17] Birth regions refer to groups of states. We use the four-region classification employed in Sepulveda et al (2006), which assigns the 32 Mexican states to North, South, Mexico City (Districto Federal) and Central. The South region is the poorest and least developed, and encompasses many of the states with high *BaseRates* (see *Figure 1*). As *Figure 1* shows, there is a regional pattern in the intensity of diarrheal mortality. Thus, *Post*Base* and *Region*Year* fixed effects will be correlated and the coefficient on *Post*Base* conditional upon *Region*Year* will provide a fairly stringent test of a trend break post-1991. In the PISA data we have only three birth cohorts and so we do not include region*year fixed effects in the math and reading equations.

crises of 1986-1988 and 1994-1996, both of which have been shown to impact infant mortality (Cutler et al 2002).

Mortality selection will tend to bias program impacts towards zero since the program will have kept alive children who otherwise may have died and these marginal survivors in the post-intervention period will tend to have relatively weak health and cognitive endowments (Almond 2006; Bozzoli et al 2009).

IV. Data

We analyze test scores for birth cohorts 1987-1993 in a window around the 1991 reform. The tests were conducted when the sample cohorts were in late childhood and adolescence, so they capture medium term impacts of the water reform at ages at which the most accelerated phase of cognitive development has been completed. We use two sources of data containing three different tests, expecting that any consistent pattern of treatment effects will show across the three tests. Further detail is in the Data Appendix. The rest of this section sets out relevant features of the data.

The Mexican Family Life Survey (*MxFLS*) is a nationally representative survey covering 8,500 households in 150 communities across 16 Mexican states (Rubalcava and Teruel 2007). For reasons discussed in the Data Appendix, we use the 2001-2002 wave of the *MxFLS* in which birth cohorts 1988-1993 are aged 8-14. The *MxFLS* administered a colored Raven's progressive matrices test, a non-verbal pattern matching assessment increasing in difficulty as the individual progresses onward, for all individuals above the age of 5. 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). The *MxFLS* records the month and year of birth and the birth state of every individual, which we match to *Post* and *BaseRate* respectively. We define *Post* as 1 if the individual was born in or after April 1990, or exposed to the April 1991 reform in the first year of life.

We model individual scores in math and reading from internationally standardized school-based mathematics and reading achievement tests conducted on 15 year olds in schools by the Program for International Student Assessments (*PISA*). We utilize data from the 2003, 2006 and 2009 waves which include cohorts born in 1987/1988, 1990, and 1993, with a portion of the 1990 (those born after April) and the entirety of the 1993 cohort denoted as being in the post-intervention group.

Using two different sources of data to study cognitive performance helps guard against shortcomings of either source from determining the results. First, the *MxFLS* is a smaller dataset (roughly 2,000 boys and girls, respectively), while *PISA* surveyed over 99,000 children across the three waves. Second, the Raven test in the *MxFLS* is often viewed as a measure of "pure intelligence", relying upon pattern recognition skills, whereas *PISA* scores measure both cognition and school-based learning. Third, the *MxFLS* was administered to a representative sample of children, regardless of whether or not they were attending school. In contrast, *PISA* tests were done on 15 year olds in school so there is potential selection into the sample who take the test given that compulsory schooling is not always well-enforced. To the extent that children who survive to 9th grade (age 15) are selected on ability, our estimates of water-reform led improvements in cognitive performance may be biased. However, as we shall see below, we find reform-led changes in cognitive performance for girls and not boys and data from the 2010 census show that school dropout rates for girls were no greater than for boys, making it unlikely that selection explains our findings. Fourth, both the *MxFLS* and the *PISA* contain rich, albeit different, sets of control variables. For instance, the *MxFLS* allows us to control for indicators of household wealth and the PISA data allow us to control for school quality. Both data sets provide indicators of parental investment, which we describe and analyze in the second part of the paper.

To model the disease environment net of income and infrastructure, we acquired a range of state specific time series data including age, gender and cause-specific mortality rates. We computed the pre-intervention level of mortality (*BaseRate)* by state and gender using the under-5 diarrhea mortality rate over the period 1988-1990.[18] We used census microdata for 1960-2000 to analyze gender-based occupational segregation in the Mexican labor market. Further information on control variables and data sources is in the Data Appendix. The means of all variables in the analysis are in *Appendix Table 1*.

V. The Impact of Early Life Diarrhea Exposure on Cognitive Development and Academic Achievement

Va. Main Results

Estimates of equation (2) for the Raven and the *PISA* school-based reading and math scores are presented in *Table 3*. For each of the three test scores, we identify significant program effects for girls. A one standard deviation decrease in the pre-program child diarrhea mortality rate is associated with a roughly 0.1 standard deviation increase in test scores for girls.[19] The coefficients for boys are consistently smaller in magnitude and, in general, not significant. The difference in the coefficients for girls and boys is statistically significant for Raven and Math, the difference for Reading is large in magnitude but not significant.[20] The fact that we see the same pattern across three different tests done at different ages on two different samples of children, one set in school and the other at home, makes the evidence of gender differences in program impact compelling. It also makes it less likely than otherwise that the results are driven by

---

[18] We also ran all equations in this paper using the infant mortality rate and the results are very similar. This is not surprising given that infant mortality accounts for over 60% of under-5 mortality and the break in trend in the under-5 rate tracks the break in trend in the infant rate. We choose to report results with the under-5 rate because of the concern that infant mortality is likely underestimated in poorer areas (see Tome et al 1997).

[19] Table 3 reports the standardized impact of a move from the 75th to 25th percentile of the *BaseRate* distribution instead of the impact of a one standard deviation change in *BaseRate*.

[20] We computed this as a t-statistic on the interaction between girl and *base_diarrhea* in equations estimated on pooled data in which all right-hand side variables were interacted with gender.

omitted trends as any omitted process would have to induce a correlation between test scores and exposure to the water reform for girls but not for boys.[21]

The estimates for girls are robust to controls for household socioeconomic status and community characteristics, state-level macroeconomic characteristics and mortality from control diseases (respiratory and vaccine-preventable diseases), fixed effects specific to birth region and birth year and, in the MxFLS data, birth state specific quadratic time trends. The control, *Post*Base_RespiratoryDisease* is a stringent control that may be interpreted as a placebo. Respiratory diseases were driven by several of the same risk factors as diarrheal disease (poverty, poor nutrition, and crowding) and declining through the sample period. Robustness of our estimates to this control implies robustness to omitted trends in common risk factors and omitted interventions that may have created a break in trend in 1991. Boys show a significant improvement in math scores in the most parsimonious specification but this is dissipated upon introducing controls.[22]

---

[21] 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). Barham et al (2013) find a 0.15 s.d. gain in cognitive outcomes for boys who were exposed to a cash transfer in Nicargua at age 0-2; they focus upon boys on the grounds that the biological vulnerability of boys is greater. Their findings contrast with ours but a fuller understanding of why would require further analysis of post-intervention investments in their data.

[22] In the specification with the richest controls, the wild bootstrap-T corrected *p-value* for the coefficient for girls in the Raven test equation is 0.058 (Table 3). In a companion paper (Venkataramani and Bhalotra 2013), we also estimated standard errors with the wild cluster boostrap-T approach in a 32 state sample and found no substantive difference in the *p*-values from this compared with those from traditional White standard errors clustered at the birth state level as in the PISA score estimates in Table 3 of this paper.

In a companion paper, we used 2012 survey data to estimate equation (2), replacing test score outcomes for adolescents with years of schooling for young adults (age 18-24). We find that infant exposure to the water reform results in girls attaining higher levels of schooling, with no impacts for boys [23]

Vb. Additional Specifications and Robustness Checks

In this section we explore a number of additional specification checks, some of which are motivated and explained in the Methods section above.

Pre-existing trends and mean reversion. The test for pre-trends shows no significant difference in the pre-1991 trend in Raven scores in states with high *vs* low pre-program diarrhea mortality rates (*Appendix Table 2).* The year-specific coefficients are jointly not statistically significant and in particular, there is no significant tendency towards convergence in Raven scores across the states before the reform the way there is after.[24] A similar potential concern is that the post-reform convergence we identify reflects mean reversion. To assess this concern, we plotted the rank correlation of state diarrhea mortality rates in 1986 and 1988 and, again, in 1985 and 1990, and the ranks look fairly stable (*Appendix Figure 1*) but we nevertheless re-estimated the equations varying the measure of *BaseRate* from the average over 1988-90 to the average over 1987-90, 1989-90 and just 1990. The results are not sensitive to these variations (available upon on request).

Flexible age of exposure, program timing, other interventions: We have so far investigated the consequences of exposure to diarrhea in infancy (first twelve months of life), assuming no impact of exposure at later ages. We investigate this assumption by

---

[23] We estimate that a one standard deviation decrease in diarrhea mortality rates is associated with a 0.2 year increase in women's schooling (95% CI 0.05-0.37); see Bhalotra and Venkataramani (2013). As discussed below, treated boys experience increases in stature, but not treated girls. Since the younger cohorts are the treated ones, any continued schooling or height growth among these young people makes our estimates conservative.

[24] Of the ten year-specific coefficients, one is significant (girls born 1988) and takes a negative sign. Since we find that reform-treated girls have better test scores, this cannot explain our results.

including new terms in the equation indicating exposure in the birth month[25] and in the second year of life. We see a fairly consistent pattern, with the largest and most well-determined impacts on test scores flowing from first exposure in the birth month and the size of the impact diminishing in the age at which the individual was first exposed to the water reform (*Appendix Table 3).* The only significant response of Raven scores in this extended specification is to birth-month exposure (first month of life), which tells us that the impacts from infant exposure (first twelve months of life) recorded in *Table 3* are driven by individuals who were born after the water reform (rather than first exposed at a stage of infancy after birth). Math and reading scores exhibit a similar pattern but the gradient in age of exposure is shallower and math scores show a significant response to exposure through age 0-2 years. Using the *MxFLS* data with which this is feasible, we investigated exposure at age 2-5 years and the coefficients were considerably smaller and completely insignificant (available on request).

These results are consistent with epidemiological and biological evidence that both the incidence of diarrhea (marking treatability by the water reform or treatment intensity)[26] and the rate of cognitive development (marking elasticity of the outcome to the treatment) are greater earlier in life. They strengthen our claim that the relationship we identify is causal and that the eventual impacts on cognition (and subsequent educational investments analysed below) are stimulated by biological changes in health and cognitive development.

Given that, in these 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 in the sample period and hence act as a placebo test of getting the timing

---

[25] Children who were born in and after April 1991 are exposed in their birth month while children born in and after April 1990 are exposed in their birth year i.e. in infancy.

[26] 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.

right, and the timing fits with the identified impacts flowing from the water reform rather than some other coincident public intervention or event.[27]

Heterogeneity in impact by household socioeconomic status: This is of substantive interest and it also serves as a consistency check on our findings. Since children from less educated households have lower test scores, finding larger impacts in less educated households would imply that the water reform stimulates convergence or catch up of lower performing children, reducing educational inequality through closing gaps in early life health. The consistency check flows from the fact that diarrhea is more prevalent among less educated (poorer) families and, so, if diarrhea reduction were in fact driving the improvement in test scores then we should see larger test score gains in less educated households. Consistent with this, we find a much weaker relationship between test scores and water reform in households in which the head has secondary education or more (*Appendix Table 4, MxFLS* sample). Exploring investment responses that may drive the test score impacts (see section VIb), we find significant increases in school attendance and a significant shift of time from chores towards homework for girls, and dividing the sample by the head's education, we find these investments in girls are driven by households with less than secondary education, consistent with the heterogeneity in test score impacts (results available on request).

***Opportunidades* as a potentially overlapping intervention:** *Opportunidades* (formerly *Progresa*) is a means-tested cash transfer program that was rolled out across rural Mexican municipalities from 1997, providing families with cash conditional upon their

---

[27] We also estimated test score equations with flexible coefficients, that is, replacing *Post*BaseRate* with a vector of terms, *i.Year*BaseRate* where *i.Year* indexes every birth cohort in the sample. In the PISA samples, there are just three birth cohorts (1987/8, 1990, 1993) and the first positive and significant coefficient is for the 1993 cohort. The *MxFLS* sample has yearly birth cohorts 1987-1993 and we estimate coefficients for every half-year. The first positive and significant coefficient is for birth cohorts April to September 1990 (who are aged one when the reform starts). Cohort and age move together in these data but these results are consistent with changes in test-scores being driven by an intervention in April 1991.

children attending schools and health clinics. The water reform 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 1987-1990), will also have been exposed to *Opportunidades,* but later in their school career, at age 8-10 and access to *Opportunidades* in the initial years may be more important for cognitive performance than access later on, for instance because children who do not start school at six may not start later even if the intervention arrives. As it seems plausible that availability of *Opportunidades* at the time of school entry may have influenced both selection into schooling and performance in school, we consider here the possibility that it was exposure to *Opportunidades* rather than exposure to the water reform that delivers the cognitive gains that we identify for girls born after April 1990.

First, if we were capturing the impact of *Opportunidades* then these impacts would be restricted to the rural sample since *Opportunidades* was, for our sample cohorts, only available in rural areas. But we find similar and not statistically different coefficients for urban and rural girls (and boys, for whom the urban and rural coefficients for both insignificant (*Appendix Table 5).* We nevertheless directly controlled for municipality-level variation in *Opportunidades* using the sample of rural households in which the head had less than secondary education, the latter being a crude proxy for the program being means-tested. We merged a measure of the intensity of *Opportunidades* coverage at the municipality level in the *MxFLS* sample (which includes municipality of residence), defining coverage as the fraction of households accessing the programme when the sample child was 6 years old.[28] Controlling for *Opportunidades* coverage at the start of primary school creates no significant change in the coefficient on the water-treatment indicator (*Post*Base*). We cannot investigate an interaction between the two programmes because *Opportunidades* was only available (at age 6 and above) to post-water-reform

---

[28] This is the number of households covered by *Opportunidades* divided by the number of total households for each municipality*year. The mean (s.d.) of this variable are 0.31 (0.24). We are grateful to Tania Barham for providing these data. We cannot match these data to PISA test scores because municipality is not identified in the PISA data.

cohorts or, equivalently, all water-treated cohorts were exposed to *Opportunidades* by age six.

Compositional effects: By virtue of lowering the price of child quality, the availability of clean water may have led to changes in fertility, and the fertility response may be heterogeneous by characteristics such as the mother's age or education (Aaronson et al 2012, Bhalotra et al 2012). This would imply endogeneity in the composition of births in the sample which may bias our estimates. For instance if, following the reform, less educated women cut back more on fertility, then the post-reform sample of births will be predisposed to higher cognitive attainment levels even in the absence of any impacts of the water reform on health and cognitive development of infants. However this sort of bias would lead us to estimate cognitive improvements for boys *and* girls born after the water reform (the sex of a birth is revealed after fertility choices are made) and since we find gains unique to girls, it is unlikely this drives our results. We nevertheless investigated this by estimating a hazard model for fertility (probability of a further birth conditional upon duration since previous birth).[29] We found a decline in fertility in response to the water reform, but no significant difference between women with secondary or lower education (results available on request). We also confirmed that there was no significant change in the distribution of the age of mothers at birth (see *Appendix Figure 2*).

Alternative mechanisms: We interpret our estimates as stemming from the impact of diarrhea infection in infancy which may directly impair brain development by diminishing net nutrition and, in addition, dampen the size and productivity of subsequent parental investments. 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 test scores either through productivity-led improvements in household income or through improved maternal health leading to better fetal and

---

[29] We used fertility histories in (earlier waves of) the ENSANUT data (see the data appendix) and "expanded" the file to allow an observation for each year of their reproductive life for every woman. The event of birth is coded as a binary variable across years within woman.

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,[30] but our finding that effects fade after infancy and are hardly discernible after the age of two is consistent with what is initially a biological mechanism. To examine the possibility that fetal health improvements (via improved maternal health or income) are a pathway, we regressed a measure of birth size from the MxFLS (a dummy variable = 1 if parents retrospectively reported their child to be "small" at birth) on the program indicator *Post*BaseRate,* now defining *Post* to indicate cohorts *in utero* during and after the reform. We find no association between birth size and exposure to the clean water program for boys or girls (*Appendix Table 6).*[31]

Another mechanism that may appear to be at play is that girls spent less time and resources collecting clean water after the water reform, and were able to invest those resources in their education. There are two reasons this cannot explain our findings. First, the time allocation data in the *MxFLS* reject the common presumption that water collection is a female-intensive activity. Among both pre- and post-program cohorts of children, boys spent more time than girls fetching water. Second, we find that girls first exposed to the water reform as infants have higher cognitive attainments as adolescents and children exposed at older ages are implicitly in a control group. This does not line up with a water fetching story as there is no reason to expect program generated

---

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

[31] The birth size question was asked for the last two births for one randomly selected woman of reproductive age in each household. The sample available for this regression is therefore smaller than the sample used in the main analysis. The sample is also unrepresentative since there are missing values for birth size and these are disproportionately greater in households headed by individuals with less than secondary schooling. This said, we have seen that it is households with less than secondary schooling that drive the impacts seen in the full sample.

discontinuities across age in time spent collecting water fifteen years after the program was put into place.[32]

VI. Explaining Gender Differences in Program Impacts

In this second section of the paper, we investigate the mechanisms driving the result that test score gains from reduced diarrhea in infancy are larger, indeed, only statistically significant for girls. We first lay out and marshall evidence in favor of the explanation we support. Then, in section VId, we consider competing explanations.

VIa. Comparative Advantage and Parental Investments – Theoretical Framework

In this section we set out the hypothesis that the stronger program impacts on cognition that we identify for girls arise from reinforcing parental investments in education that favor girls. This is consistent with girls having a comparative advantage in skill- (brain-) intensive tasks relative to boys, who hold a comparative advantage in tasks involving physical strength (brawn),[33] which generates gender-based occupational sorting, with women sorting differentially into brain-intensive occupations. As a result, the average return to educational investments is higher for girls, and this feeds back into optimal parental investments in education favoring girls, while optimal investments in brawn favor boys.

Schooling and nutrition are complementary with the individual health endowment in the production function for human capital, so an improvement in the health endowment raises the productivity of any subsequent investments, stimulating increased

---

[32] It is plausible that girls with healthier infancies were less likely to fetch water later in childhood and we do find a negative association between water collection times and *Post*Base_Diarrhea* in the *MxFLS* (available upon request) but this is for both girls and boys and the coefficient is larger for boys.

[33] The biological premise for this is that men are endowed with more brawn than women, as is evident for example in grip strength (Mathiowetz et al 1985; Günther et al 2008), and nutritional inputs that augment body mass tend to increase brawn substantially more for males than for females (e.g., Round et al, 1999).

investment. Gendered comparative advantage results in gender differentiation of these investments.

Recent work formalizes these predictions (Pitt et al 2012; Rendall 2010; Rosenzweig and Zhang 2012).[34, 35] Here we focus on relevant testable predictions of the model. The water reform led to an improvement in infant health which the model predicts will lead to increased investment in cognitive skills for girls. Whether there is an increase in educational investments in boys is ambiguous, indeed, educational investments may decrease.[36] Importantly, the positive relationship between endowment quality and educational investments grows stronger as skilled opportunities in the economy increase. The expansion of skilled jobs (the relative decline of jobs that require physical strength) also stimulates an increase in the labor force participation of women. Incorporating responses at the extensive margin makes the model predictions starker (see, for instance, Rendall (2010), who shows that skill-biased technical change in the US economy in the post-war era can explain a substantial share of the rise in women's labor force participation and the closing of the gender wage and education gaps).

Pitt et al (2012) show that improved health endowments in Bangladesh led to increased investments in brawn and less schooling for boys alongside small positive changes in schooling investments and outcomes for women. Using data from the more dynamic and skill-intensive economy of China, Rosenzweig and Zhang (2012) find that

---

[34] Galor and Weil (1996) also posited a comparative advantage of women in brain-intensive occupations in their analysis of fertility and economic growth.

[35] A Roy-type model of sorting is specified in which workers are bundles of brain (cognitive skill or academic achievement) and brawn and they sort among occupations according to their comparative advantage in these two attributes (Roy 1951). Parents maximize their utility which depends upon the child's adult wage and current consumption, subject to budget and technology constraints. Family income and wages from child labor contribute income, and schooling and consumption incur costs. For men (only), brawn is increasing (with diminishing returns) in the infant health endowment. The wage function for youth (child labor) is a function of brawn alone while the adult wage function depends on brawn and skills. There is dynamic complementarity among inputs to the production of human capital.

[36] For boys, improved early childhood health additionally increases returns to brawn and thereby increases the opportunity cost of schooling (given that youth labor is a function of brawn only).

increased birth weight (identified within same-sex twin pairs) is associated with large increases in schooling, and test scores for girls. We contribute to this line of evidence using data from Mexico which is more skill-intensive than China. As discussed in the introduction, this study differentiates itself from these previous two investigations of the brain-brawn hypothesis by its ability to model a policy-driven quasi-experimental change in the infant endowment. Also, we fill out the picture, directly modeling the impact of an exogenous change in the infant endowment on both parental investments in children and their cognitive attainments, alongside analysis of trends in labor force participation and occupational sorting, allowing for education gradients.

So as to establish the relevance of occupational sorting in interpretation of our results, we first investigate gender-based occupational sorting, and trends in schooling and labor force participation for women (relative to men) as the skill level of the economy increases. We then use the microdata samples analyzed for cognitive performance, to investigate whether parental investments in skill formation reinforce the positive endowment shock flowing from access to clean water in infancy and how these investments are differentiated by gender. The results are presented in the following two subsections.

VIb. Evidence: Parental Investment Responses

To examine whether parents' education-related investments increased more for girls than boys as a result of exposure to the National Clean Water Program, we re-estimate equation (2), replacing test scores with indicators of investments available in the *MxFLS* and *PISA* surveys. From the *MxFLS*, we have information on child time use (time spent on homework, household labor or chores) and school attendance. The PISA survey includes information on whether the child attended preschool and on educational items a child has available at home.[37] Summary statistics of these variables are in *Appendix Table 1*.

The results are in *Table 4*. For girls, a 1 standard deviation decrease in diarrheal mortality flowing from the water reform is estimated to have resulted in an increase in preschool attendance of 3.02 percentage points and of school attendance by 1.38 percentage points. The time spent on homework increased by roughly 41 minutes and the index of parental investments in educated-related infrastructure at home increased by 11.6 percentage points. The estimated effects for boys are far smaller in magnitude and not statistically significant from zero, except for school attendance, which shows a positive coefficient but of smaller magnitude than that for girls.[38] These results are robust to the inclusion of the controls discussed in Section III (we only present estimates from models with the full set of controls here).[39] If, as recent research suggests, investments enter multiplicatively in the production of cognitive attainment then small changes in the identified investments may have large effects.[40]

---

[37] This includes desks, a quiet place to study, a personal computer, educational software, internet access, literature, poetry, textbooks, and a dictionary. We sum these binary variables to create an index of the number of such items.

[38] School attendance rates at age 8-14 (the age of our sample cohorts in MxFLS 2002) are high (see Appendix Table 1) and water-treated girls and boys may show improved attendance rates not only because parents invest in terms of keeping them in school but because they are less often off sick. Using a 2012 survey in which the sample cohorts are aged 18-24, we find a significant increase in years of schooling completed for water-treated girls and no increase for boys.

[39] The t-statistic on the homework coefficient for girls is 1.73, so it is on the margin of significance. If we re-estimate this model controlling for birth-month exposure and exposure in the second year of life then the birth-month and infant exposure coefficients are both statistically significant (Appendix Table A7).

[40] Several studies suggest causal relationships between the particular investments we study and cognitive development. For preschool, see Apps et al (2012), Behrman et al (2004); for homework (in fact home teaching) see Behrman et al (1999). We estimated conditional and unconditional correlations of the parental investment inputs in *Table 5* with Raven and *PISA* scores and found positive and significant relationships. Preschool attendance raises reading and math scores at age 15 by 6% mean (0.3 s.d.), an additional hour of homework in late childhood raises Raven scores by 0.5% mean (0.014 s.d.) and school attendance raises Raven scores by 17% mean (0.45 s.d.). These are of course only correlations since these investments are all endogenous.

VIc. Evidence: Female Schooling, Labor Force Participation, and the Gender Division of Labor in the Mexican Economy

The Mexican economy has changed dramatically since the mid-20th century, having become substantially more skill-intensive. In order to study employment and occupation trends by gender, we assimilated census micro-data for 1960 to 2000 from IPUMS. So as to track calendar time trends for adult cohorts (as distinct from cohort trends), we computed statistics for individuals aged 20-50 years at enumeration. Based on a newly available linkage of Mexican occupational categories to job characteristics from the US Dictionary of Occupational Titles (Vogl 2012), we classified individuals in employment as being in brain- *vs* brawn-intensive occupations.

We estimate labor force participation and gender-based occupational sorting by estimating the equations displayed in *Table 5* for each available census year in 1960-2000 (no census was conducted in 1980), so that looking across columns one can see the trend in the coefficients. The ratio of female to male employment rose from about a fifth to a half between 1960 and 2000 and increases in labor force participation were greater amongst more educated women (panel A). There is evidence of sorting: conditional on employment, women were 10 to 20 percentage points more likely to be in brain-intensive occupations than men and educated women more so. Again there is a positive trend (panel B).[41] Since education is endogenous, the estimates in the right hand column where interactions of gender and education are presented are only descriptive. Overall the evidence in Table 4 shows that women, particularly educated women, were increasingly entering the labor force and sorting into brain-intensive jobs, which suggests that parents making investments in our sample cohorts were witnessing increased returns to educational investments in girls more than in boys. Our evidence is consistent with

---

[41] The share of women in brain-intensive occupations shows an upward blip in the 1990 census. We exaxmined the data but could not identify any reason that this is not genuine. We rest nothing upon this coefficient so here we simply flag it.

findings from other work. For example, a recent study by Aguayo-Tellez et al (2010) illustrates how trade liberalization as a result of the North American Free Trade Agreement (NAFTA) led to increased employment opportunities and wages for women.[42]

Figures 5 and 6 show related plots. *Figure 5a* plots the share of brawn-intensive industries (agriculture, construction, and mining) as a function of GDP between 1970 and 2000 using National Accounts data. We see a large increase (decrease) in the skill-intensity (brawniness) of the economy, which is particularly pronounced starting in the 1980s, when Mexico liberalized foreign trade and investment policies. This was reinforced in 1994 with the NAFTA agreement (Hanson 2003). *Figure 5b* shows, using the cross-section of Mexican states, that the share of brain-intensive occupations increases as the share of agriculture, construction and mining increases. Concomitant with the trends in skill intensity in *Figure 5a*, census data show that women are over-represented in brain-intensive occupations and that this tendency is growing stronger over time (*Figures 6A, 6B*). Among men, even in the most recent census data, more than 80% work in brawn-intensive occupations. Thus, the Mexican economy, while characterized by increasing skill-intensity (and, therefore, more labor force opportunities for women) continues to have a significant brawn-intensity. The proportion of women in the labor force has shown a rapid rise (from 10% in 1960 to nearly 40% in 2000, *Figure 6C*), and women's schooling (*Figure 6D, 6E*). has also increased more rapidly for women than for men.

VId. Differential Program Impacts, Selection, Biological differences

---

[42] Recall that we use the pre-intervention level of diarrhea in the birth state to denote intensity of the intervention (size of reduction in diarrhea following the water reform). If states with high pre-reform diarrhea were systematically "brawnier" states (for instance because diarrhea is correlated with poverty) then we may be capturing impacts of the occupational structure on parental investments together with impacts of the treatability of the state by a disease control intervention. This would only exacerbate the gender difference but it is of some concern that it may, on average (i.e. across gender), bias our interpretation of coefficients. *Appendix Figure 4* allays this concern, showing that there is no systematic relationship between pre-diarrhea and the share of brawn-based occupations in the state.

In this section we consider other explanations of the identified gender differences in outcomes and investments. A natural explanation of our finding that improved cognitive performance flowing from exposure to clean water during infancy is larger for girls than boys is that diarrhea among girls fell more sharply as a result of the reform. However, estimates of the "first stage" show that the intervention was associated with a decline in diarrhea for boys that was no smaller (indeed slightly larger) than the decline for girls (*Tables 1* and *2)*.

A second possibility is greater survival selection among boys. Following the implementation of the clean water intervention, the marginal survivor will have been relatively frail and, given their higher biological vulnerability in early childhood (Kraemer 2000), one might expect greater (negative) selection among boys. If innate ability is lower amongst children in the left tail of the health distribution then it is possible that our finding of weak test score impacts for boys arises from selection. However, diarrhea mortality rates for boys were only slightly larger than for girls and, as discussed, the reform-led decline in diarrhea (in absolute and proportional terms) was similar for boys and girls, so selection effects are unlikely to be large enough to explain our results. The role of selection is also undermined by the finding that boys exposed to the program experienced gains in height and these were concentrated at the lower end of the height distribution, where selection effects should be strongest (Venkataramani 2009; Venkataramani and Bhalotra 2013).

Another form of selection but one that is only germane to the *PISA* test sample is differential survival in school to age 15. If girls were more likely to have dropped out by age 15, and the girls who persist in school are of selectively high ability, this could account for our finding that program effects are larger for girls. However, using data from the 2010 census (in which the marginal cohort in our sample, born in 1991, is 19) we confirm that for our sample cohorts the dropout rate by age 15 is small, similar across the sexes and if at all slightly larger for boys. This is confirmed in looking at the gender ratio of the Mexican PISA test sample, in which 47 percent are boys, consistent with slightly higher dropout among boys. The net secondary school enrollment rate in Mexico is 74% for girls and 71% for boys and survival rates to grade 5 are 94% for girls and 91% for boys (UNESCO 2010). Another reason that selection into school is unlikely to explain the better performance of girls in the PISA tests is that girls also do better on Raven tests which were administered unconditional on school attendance A similar gender pattern emerges in our analysis of parental investments in education using *MxFLS* data that, again, pertain to all children and not those who select to stay on in school to age 15.

A fourth possibility is that the differential results reflect catch-up if girls were initially lagging boys in educational investments and test scores.[43] However, although boys did have an initial advantage in math scores, Raven and reading scores were very similar between the sexes prior to the program, as were the range of educational investments that we consider below. Finally, might our results arise from brain development in girls being more sensitive to early life health shocks than boys? There is no biomedical evidence to support this, if anything, boys' brains may be more sensitive to *in utero* and early life insults (Kraemer 2000).

IVe. Summary

Overall, the results in part-2 of the paper are consistent with the predictions generated by models of human capital investment predicated on gender differences in comparative advantage in brawn *vs* skills. Parental investments in educational inputs increase with reduced diarrhea in infancy and much more so for girls. Consistent with the theory, treated boys appear to have received increased investments in brawn, indicated by their height, while women experienced no increase in height (Venkataramani and Bhalotra 2013). Macroeconomic trends in the Mexican economy indicate higher relative rates of

---

[43] Assuming a production function for cognition that shows diminishing returns to educational inputs, if girls were initially at a more concave point on the curve then the marginal returns to investments would be larger for girls.

return to girls' schooling and boys' brawn, which are consistent with the observed gender-differentiation in parental investments.

Since we show that the reform-led drop in diarrhea is similar for boys and girls, and that test score improvements are nevertheless unique to girls, it follows that reform-induced parental investments in girls' education are critical to the realization of impacts of diarrhea reduction in infancy on test performance. Our results in this way, implicitly provide evidence of dynamic complementarities across inputs in the production of cognitive skills.

We now discuss supporting evidence. First, the parental investments that we examine are known to have independent strong impacts on cognitive skills (see footnote 40), which supports the plausibility of our contention that the investments are critical. Second, our finding that the direct effect of the clean water reform (or, equivalently, the infant health shock) was no different for boys than for girls effectively addresses the critique by Bleakley (2010) that, at optimal levels of investments, the marginal return to additional investments on the outcome of interest is zero (by the Envelope Theorem) and so only the direct effect of endowments on the outcomes is substantively important. For this argument to work here, we would need to see positive impacts for boys' cognitive outcomes, which we do not.[44] Our argument becomes all the more potent in view of our finding large impacts on height for treated male cohorts alongside no impacts for females (Venkataramani and Bhalotra 2013). The mapping of impacts on cognitive and physical outcomes by gender in a manner consistent with sex differences in rates of return to these attributes in the labor market supports the notion that the "phenotype" of adult capacity, be it brains or brawn, depends on the nature of subsequent parental investments.

---

[44] Alternatively, the direct impacts of a positive health shock on cognitive capacity would have be larger for girls than for boys. There is no known biological reason why this would be, if anything, the biological literature suggests that boys' cognitive development should be more sensitive to health shocks (Kraemer 2000).

Our results support the findings of earlier work in which we analyze the long run socioeconomic impacts of a sharp drop in infant pneumonia driven by the introduction of antibiotics (Bhalotra and Venkataramani 2012). We find large long run gains for white men and limited long run gains for black men. The drop in infant pneumonia was no larger (in fact was smaller) for white children, ruling out differential short term program effects as an explanation of the race differential in long run outcomes. Using state variation in indicators of racial segregation, we demonstrate that the outcome gains were weaker for blacks because reinforcing parental investments were muted for them on account of segregation-related caps on returns to investment. As here, identification of endowments exploits a health intervention and identification of subsequent human capital investments exploits institutionally driven group differences in rates of return to investment. In both cases, the group for which returns to investment are lower witnesses much weaker long run gains despite similar reductions in disease in infancy. This demonstrates the empirical importance of complementary investments.

Our work relates to a recurrent, more generic problem in the wider literature which is that even if we can identify changes in investments, we typically cannot identify the weight of investments in delivering outcomes because the investments are endogenous in the outcome equation. For instance, Carneiro et. al. (2012) use instruments to identify the impact of maternal education on investments and outcomes amongst American children, but are unable to identify the weight of investments in determining the outcomes.

Our argument for dynamic complementarity is also supported by the analysis in Venkataramani (2012) who shows that early life exposure to malaria eradication for cohorts born in the late 1950s in Mexico was associated with improved cognition for men, but not women. As there is no biological reason (or evidence) for malaria and diarrhea to have different impacts on male and female infants, it seems likely that it is differences in responsive investments that drive differences in outcomes in this vs the cited study, and Venkataramani (2012) reports evidence of investment differences consistent with the outcome differences by gender. A plausible explanation of the gender reveral in results is that women's labour force participation rates in the 1950s were very low (10%, Figure 6E).[45]

VII. Conclusions

A growing body of work highlights the importance of cognitive skills in driving socioeconomic outcomes, making it important from a public policy standpoint to understand how to help individuals achieve their cognitive potential and, related, to narrow gaps in cognitive skills. We find that infectious disease has a causal influence on cognitive performance, and this appears to depend critically on diminished infant endowments depressing subsequent human capital investments in children. We establish that parental investment responses are larger in states in which the reduction in diarrhea is greater and these are, systematically, states in which the pre-intervention incidence of diarrhea was greater. This implies that interventions such as clean water provision, acting through infant exposure, can contribute to narrowing the documented cross-country gaps in cognitive performance.

Our findings also provide evidence to support the somewhat novel notion that a substantial share of the gender gap in test performance[46] within and across countries may be explained by the prevalence of diarrhea (and possibly other infectious diseases or other determinants of infant health and cognitive endowments). Although there is, in our sample, no significant gender difference in the endowment-diminishing impact of diarrhea

---

[45] An explanation for cognitive impacts for men, who hold a comparative advantage in brawn, is that, as in the Pitt et al and Rosenzweig and Zhang models, there are complementarities between skills and brawn in determining productivity and wages. In contemporary Mexico, we perhaps do not find any impact on boys' cognition because the margin for which skill accumulation matters for brawny labor is likely met with compulsory schooling through to the age of 15 (with which compliance is not perfect but high: see Appendix Table 1).

[46] Girls tend to do less well in math across countries. However their reading scores are not systematically worse than those of boys and, often, are better. Our results imply that any reading advantage that girls in countries with high diarrhea prevalence display would be even greater in the absence of diarrhea infections in infancy. So the spirit of our argument holds as what is desirable is not to close gaps per se but to allow every individual to approach their potential..

infections in infancy, there are significant differences in parental responses, which are consistent with the relative return to investing in education for girls versus boys. Interestingly, this too, stems from biology, which endows boys with a comparative advantage in brawny tasks or, equivalently, a comparative disadvantage in non-brawny or skilled tasks. So the potential for clean water programs (or possibly other infant health interventions) to narrow cognitive performance gaps across gender would appear to depend upon the demand for skill in the economy and the extent of gender-based occupational sorting, *viz* the relative returns to educational investments in girls *vs* boys.

At the time of writing, our sample cohorts are 19-25 years old, they have not all matured onto the labor market and some are still in education. In a few years, when they are more established in the labor market, it will be useful to analyze labor force participation, occupational choice and wages by gender. We will then be able to capture the impacts of an early life intervention such as a clean water reform on the evolution of female wages and, in particular, the trajectory of the female-male wage gap.

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. 2006. "Is the 1918 Influenza Pandemic Over? Long-Term Effects of *in Utero* Influenza Exposure in the Post-1940 U.S. Population." *Journal of Political Economy*, 114(4), 671-712.

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,

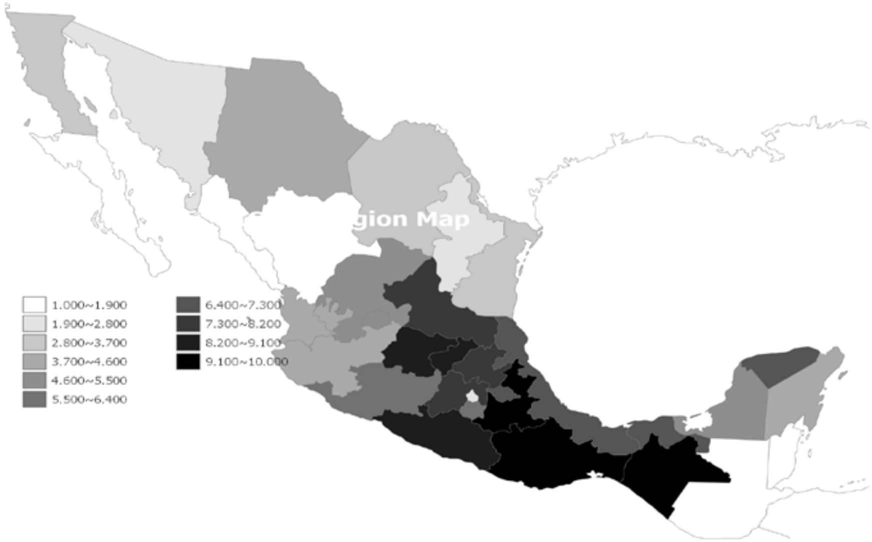
Venkataramani, Atheendar and Sonia Bhalotra. 2013. "Effect of Early Life Exposure to a Clean Water Intervention on Adult Stature and Educational Attainment in Mexico: A Quasi-Experimental Retrospective Cohort Study." *Mimeo, University of Essex.*

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.

Figures

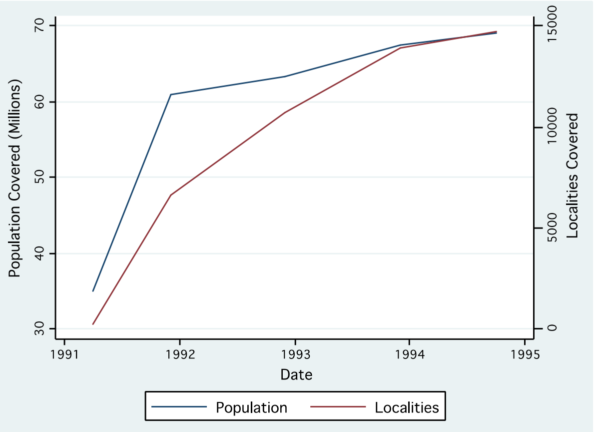
Figure 1 – Child Diarrheal Mortality Rates Across Mexican States, 1988-1990 (Pre-Intervention Baseline)



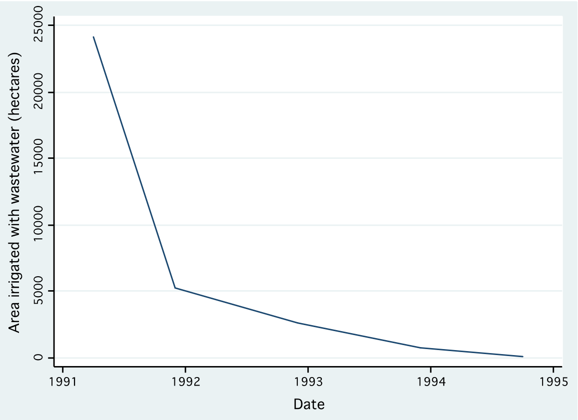
Source: Mortality data from the Mexican Secretary of Health (*Secretaría de Salud*). Population estimates for the deflator from the National Council on Population (*Consejo Nacional de Población*). Map drawn by authors.

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

A. Population Access to Chlorinated Water

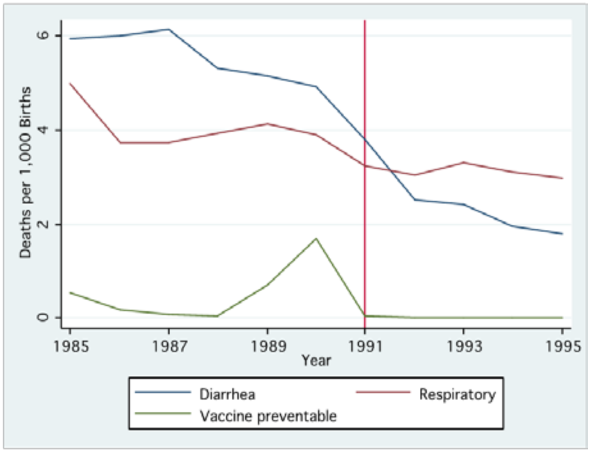


B. Land Area Irrigated with Waste Water



Source: Government of Mexico, National Water Commission

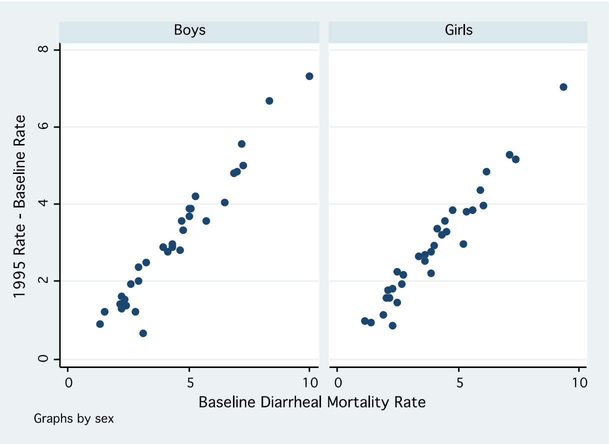
Figure 3 – Trends in Child Infectious Disease Mortality: Diarrhea, Respiratory Infections, and Vaccine Preventable Infections



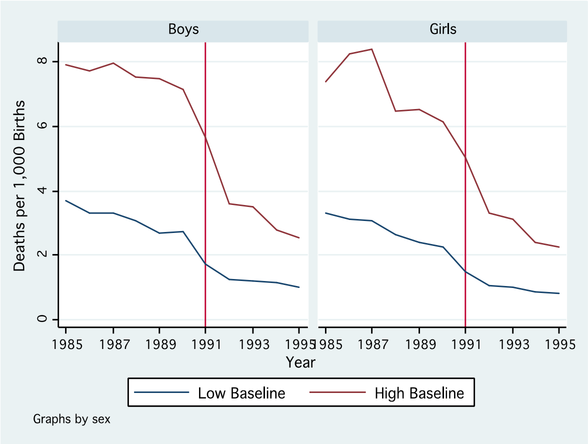
Source: See *Figure 1*.

Figure 4 – Convergence across States in Child Diarrhea Mortality Rates

A. Scatter Plot of Absolute Change in Diarrheal Mortality Post-1991 against thePre-Intervention Rate, State Means

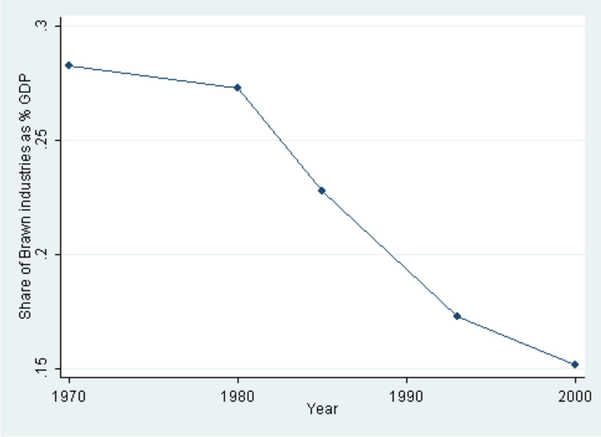


B. Trends by Gender for Pre-Program Diarrhea Mortality Rates Below and Above the Median Rate



Source: See *Figure 1*.

Figure 5a – Trend Decline in Share of Agricultural, Mining, and Construction as % of GDP



Source: National Accounts Data provided by Ernesto Aguayo-Tellez.

Figure 5b- Share of Employment in Brain-Intensive Occupations is Decreasing in Agriculture, Mining and Construction Share : State Means

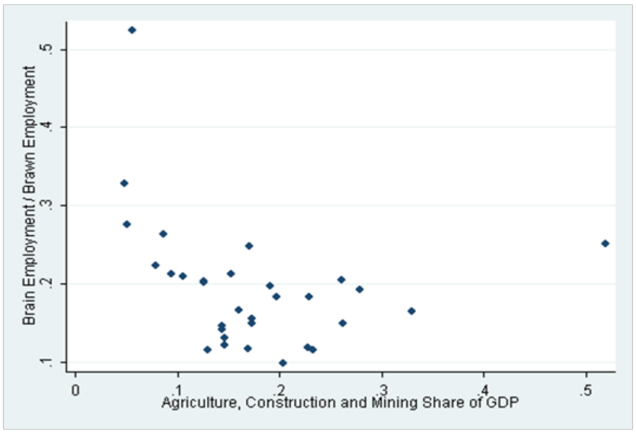
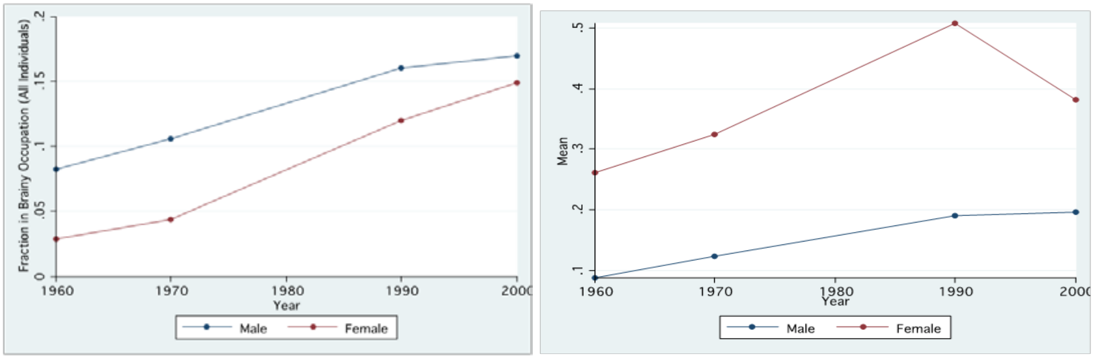
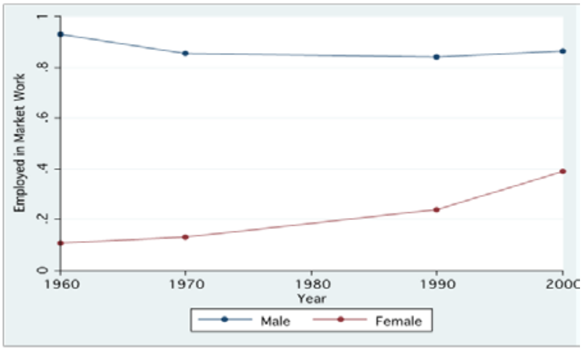


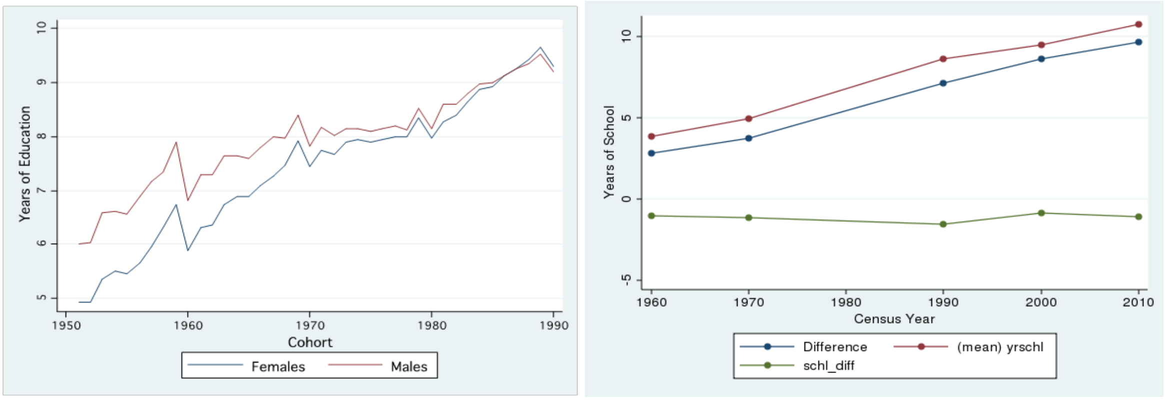
Figure 6 – Trends in Occupational Sorting, Employment and Schooling by Gender



(a) The left panel above plots the fraction of all individuals in brain-intensive occupations, showing women (the lower line) catching up with men by virtue of a stronger trend (B) the right panel plots this same fraction for individuals in employment, showing the fraction for women (the upper line) is consistently higher than for men. Source: Census microdata



(C) The left panel above shows a stronger trend in employment for women (the lower line) than for men



(D) the left panel shows the years of education of women (lower line) are converging towards those of men (MxFLS data) and (E) the right panel shows that in the employed sample, women (the upper line) have more education than men. The lowest line is the difference between years of school of employed men and women. Source, unless otherwise stated: Census microdata

Table 1 – "First Stage": Trend Break in Child Diarrheal Disease Mortality Following the National Clean Water Program

	(1)	(2)	(3)
	Pooled	Girls	Boys
Post	-0.381***	-0.346***	-0.417***
	(0.0416)	(0.0549)	(0.0518)
Year	-0.0638***	-0.0786***	-0.0490***
	(0.0126)	(0.0109)	(0.0161)
Post*Year	-0.0994***	-0.0934***	-0.105***
	(0.0183)	(0.0217)	(0.0207)
N	704	352	352

Note: Robust standard errors, clustered at the state level, in parenthesis. *** - p<0.01. Dependent variable is the logarithm of the diarrhea mortality rate for children under the age of five. Year is set to zero for 1991 and Post is set to 1 if Year>=0.1. The sample includes the 32 Mexican states over the period 1985-1995. "Pooled" refers to pooling data for boys and girls. Every equation includes state and year fixed effects. See the text for details.

Table 2 – "First Stage": Trend Break in Diarrheal Following the Clean Water Program: Impact of National Clean Water Program on Child Diarrheal Mortality, Differences-in-Difference Estimates s with Respiratory Disease Mortality as a Control

	(1)	(2)	(3)
	Pooled	Girls	Boys
Post	-0.503***	-0.580***	-0.427***
	(0.0836)	(0.123)	(0.0760)
Year	-0.0165	-0.0225	-0.0105
	(0.0134)	(0.0136)	(0.0145)
Post*Year	0.0121	0.0353	-0.0111
	(0.0195)	(0.0365)	(0.0141)
Treated*Post	0.0956	0.159	0.0324
	(0.127)	(0.195)	(0.114)
Treated*Year	-0.0731***	-0.0775***	-0.0688***
	(0.0174)	(0.0176)	(0.0185)
Treated*Post*Year	-0.0857***	-0.107**	-0.0641***
	(0.0257)	(0.0498)	(0.0141)
N	1408	704	704

Note: Robust standard errors, clustered at the state level, in parenthesis. *** - p<0.01. Dependent variable is the logarithm of the mortality rate. Treated = 1 for diarrhea and 0 for respiratory disease. See Notes to Table 1. The coefficient on the triple interaction term indicates the difference in any trend break in child mortality rates from diarrhea relative to respiratory diseases. The equations include state, year, and disease fixed effects.

Table 3 – Impacts of Clean Water Program Exposure in Infancy on Raven, Math and Reading Test Scores

	Girls				Boys			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Raven Scores	0.378*	0.569**	0.767**	0.857**	-0.170	-0.399	-0.327	0.0151
	(0.202)	(0.260)	(0.293)	(0.296)	(0.197)	(0.346)	(0.299)	(0.269)
Wild cluster-t p-value	*0.198*	*0.158*	*0.104*	*0.058*	*0.236*	*0.272*	*0.308*	*0.603*
Effect size s.d	0.058	0.087	0.118	0.132	-0.028	-0.066	-0.054	0.002
N	2,527	2,527	2,527	2,527	2,435	2,435	2,435	2,435
Reading Scores	2.871***	1.868*	1.880*		1.648	0.0474	-0.237	
	(0.969)	(1.029)	(0.978)		(1.119)	(1.248)	(1.231)	
Effect size s.d	0.110	0.069	0.070		0.059	0.002	-0.013	
N	52,991	52,991	52,991		46,211	46,211	46,211	
Mathematics Scores	3.294***	2.419***	2.366***		2.128**	0.640	0.416	
	(0.839)	(0.843)	(0.816)		(1.007)	(1.059)	(1.064)	
Effect size s.d	0.110	0.076	0.074		0.072	0.022	0.014	
N	52,991	52,991	52,991		46,211	46,211	46,211	46,211
Controls								
Birth State FE, Birth Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Household and School Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State*Year Controls	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Birth Region X Birth Year FE	No	No	Yes	Yes	No	No	Yes	Yes
Birth State Quadratic Trends	No	No	No	Yes	No	No	No	Yes

Notes: Robust standard errors clustered at birth state level in parentheses. *** - p<0.01, ** - p<0.05, * - p<0.10. For the *MxFLS* which has 16 rather than 32 states, we also report the p-value from the wild cluster-T bootstrap method (Cameron et.al. 2008). The sample comprises individuals born between 1987 and 1993 (inclusive). Raven Progressive Matrix test scores (MxFLS 2002) are defined as fraction of questions answered correctly, range 0 to 100. Reading and Mathematics scores from *PISA* surveys, 2003, 2006, 2009, are available normalized on an international scale. Each cell of the Table reports the coefficient on *Post*BaseRate* (equation 2 in main text) for the three different tests (Rows) using different sets of controls (Columns). *BaseRate* is the gender-specific child diarrheal mortality rate in the birth state averaged across the pre-intervention years, 1988-90 and *Post*=1 if child is born on or after April 1990 (i.e. exposed to program in the first 12 months of life). For the Raven test sample, "*Household and School*" controls include household head's level of education and gender, housing characteristics (type of wall, floor), logarithm of consumption expenditures and the size of the community the child lives in. "*State*Year*" controls include *Post* interacted with each of the pre-intervention birth state child mortality rate from respiratory diseases, mortality from vaccine preventable diseases, log GDP per capita, literacy and average grade attainment. For the Reading and Mathematics samples, the controls are similar except that household controls are now comprised of indicators for maternal and paternal schooling attainment, occupational qualification, school size, measures of available educational materials, whether school is public or private, and whether the school is situated in an urban or rural area. The effect size s.d. is calculated as the standardized impact of a 1 s.d. decrease in *BaseRate (Diarrhea)*. We tested for the significance of the gender difference in the coefficients in the specifications with the richest controls by estimating a pooled model in which all right hand side variables are interacted with gender. The p-values for Raven, Reading and Math scores are 0.00, 0.27, 0.10 respectively.

Table 4 - Impacts of Clean Water Program Exposure in Infancy on Parental Investments

	Girls	Boys
Attend School	0.00488*	0.00401**
	(0.00276)	(0.00158)
Time on Homework	0.144	-0.0355
	(0.0828)	(0.0687)
Attend Preschool	0.0116***	0.00228
	(0.00199)	(0.00488)
Parental Investment Index	0.0409**	0.0202
	(0.0151)	(0.0262)

Note: Robust standard errors clustered at birth state level in parentheses. *** - p<0.01, ** - p<0.05, * - p<0.10. Every cell reports the coefficients on *Post*BaseRate* from a separate regression, see notes to Table 3, the estimated equation is equation (2) in the paper with test scores replaced by indicators of parents' education-related investments. School attendance refers to the time of survey and time on homework is in hours per week, these variables are from MxFLS 2002. Preschool attendance and the Parental Investment Index are constructed from questions to parents concerning educational related items such as desk, computer that a child has in the household (see text). These are taken from PISA 2003, 2006 and 2009. Sample sizes are as in Table 3, controls are as in the last column in Table 3. Appendix Table A7 shows impacts of exposure in the birth month and infancy are significant for girls in a specification that controls also for exposure in the second year of life. However in such a specification, school attendance shows no significant response.

Table 5 - Employment and Occupational Sorting by Gender and Education, Mexican Census 1960-2000

Labor Force Participation

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Census Year	**1960**	**1970**	**1990**	**2000**	**1960**	**1970**	**1990**	**2000**
Female	-0.798***	-0.704***	-0.580***	-0.501***	-0.846***	-0.796***	-0.797***	-0.704***
	(0.0199)	(0.0252)	(0.0209)	(0.0166)	(0.00898)	(0.0108)	(0.0110)	(0.0120)
Education	0.0110***	0.00722***	0.0154***	0.0172***	0.00233	-0.00438***	-9.09e-05	0.00333**
	(0.00145)	(0.000902)	(0.000773)	(0.000927)	(0.00235)	(0.00141)	(0.00127)	(0.00137)
Female*Education					0.0193***	0.0272***	0.0331***	0.0269***
					(0.00241)	(0.00189)	(0.00113)	(0.00108)
N	135,920	154,610	2,873,475	3,702,327	135,920	154,610	2,873,475	3,702,327

Employment in Brain-Intensive Occupations

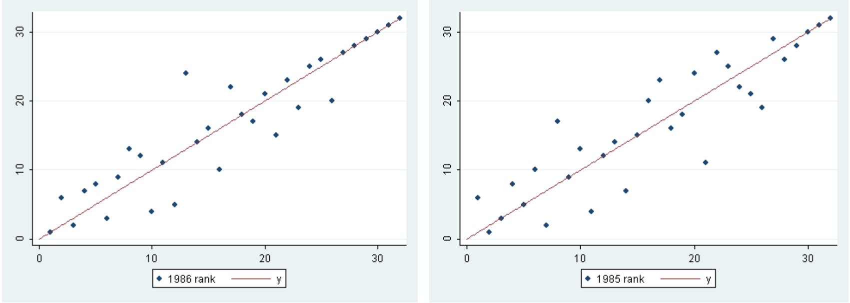
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Census Year	**1960**	**1970**	**1990**	**2000**	**1960**	**1970**	**1990**	**2000**
Female	0.106***	0.138***	0.234***	0.129***	0.0234**	0.0332***	0.0658***	-0.0493***
	(0.0106)	(0.00582)	(0.00341)	(0.00345)	(0.0102)	(0.00539)	(0.0137)	(0.00854)
Education	0.0513***	0.0508***	0.0513***	0.0533***	0.0472***	0.0465***	0.0465***	0.0463***
	(0.00332)	(0.00237)	(0.00153)	(0.00194)	(0.00346)	(0.00243)	(0.00178)	(0.00218)
Female*Education					0.0223***	0.0223***	0.0202***	0.0206***
					(0.00309)	(0.00219)	(0.00214)	(0.00129)
N	70,081	76,285	1,511,215	2,188,066	70,081	76,285	1,511,215	2,188,066

Note: Robust standard errors corrected for clustering at the state level in parenthesis. *** - p<0.01, ** - p<0.05, * - p<0.10. Census microdata, people aged from 20 to 50 years old at the time of enumeration, census year is in the column header. Dependent variables: Employment = 1 if the individual reports working, 0 otherwise. Employed in a Brain-Intensive Occupation = 1 if the individual is working in an occupation considered brain intensive as per Vogl (2012), 0 otherwise. Education is years as schooling. Every equation includes birth year, census year and enumeration state fixed effects.

Appendix Figures

Figure A1– Tests of Mean Reversion: Rank Correlation in State Diarrheal Mortality Rates

A. 1986 against 1988 rank B. 1985 against 1990 rank



Source: See *Figure 1*.

Figure A2- Distribution of Mother's Age at Birth, Pre and Post 1991.

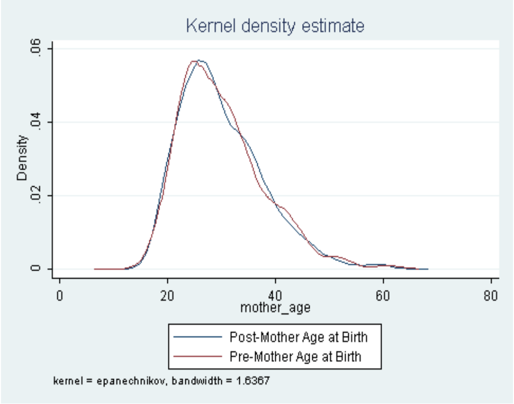
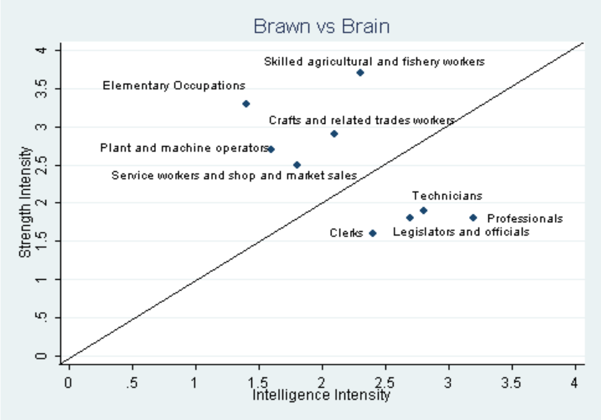
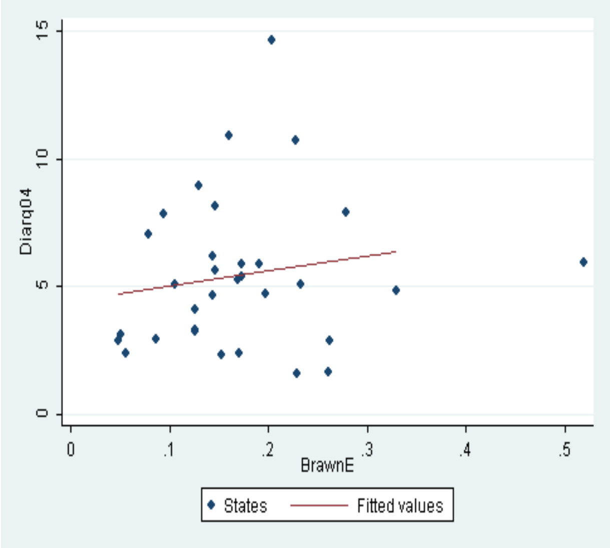


Figure A3 – Classification of Occupations into Brain and Brawn Intensive



Source:The plot depicts a classification of nine ISCO occupations which can be matched to Mexican and US job classifications. It is based upon a mapping of the Dictionary of Occupational Titles to the 1980 US census, applied to Mexico 2002 survey data (MxFLS) by Vogl (2012). We apply it to the MxFLS and census data from Mexico. The strength *vs* intelligence classification is robust to other classifications such as numerical skills *vs* physical demands of the occupation.

Figure A4 – State Scatter of Diarrhea Mortality Rates Against Brawn-Intensity



Source: Brawn intensity of a state is the share of occupations that are brawn-intensive, obtained by computing means from census microdata.

Table A1 - Descriptive Statistics

	Girls			Boys		
	Mean	S.D.	N	Mean	S.D.	N
MxFLS						
Outcomes						
Raven Scores	61.282	21.499	2,691	62.417	20.928	2,595
Attend School	0.942	0.233	2,718	0.947	0.224	2,642
Time on Homework	6.848	5.611	2,422	6.367	5.340	2,360
small birth size	0.253	0.435	1,127	0.276	0.447	1,131
Controls						
BaseRate (Diarrhea)	3.209	1.826	2,559	3.686	2.010	2,488
BaseRate (Respiratory)	0.080	0.089	2,559	0.093	0.106	2,488
BaseRate (Vaccine Preventable Disease)	5.579	3.555	2,559	5.721	3.724	2,488
Birth State Literacy Rate	0.953	0.028	2,559	0.953	0.028	2,488
Birth State Mean Schooling (25-35 year olds)	7.846	0.971	2,559	7.858	0.967	2,488
Ln(GDP) in Birth State	9.294	0.341	2,559	9.305	0.345	2,488
Ln(Rainfall) in Birth State	6.520	0.470	2,559	6.511	0.498	2,488
Household Head Education (1-3 scale)	1.238	0.663	2,731	1.256	0.668	2,654
Female HH Head (=1)	0.172	0.378	2,731	0.170	0.376	2,654
Community Size	2.749	1.342	2,729	2.745	1.338	2,653
Floor Made of Cement, Wood (=1)	0.828	0.378	2,731	0.841	0.365	2,654
Wall Made of Concrete, Wood (=1)	0.750	0.433	2,731	0.749	0.433	2,654
Roof Made of Aluminum, Shingles (=1)	0.616	0.486	2,731	0.629	0.483	2,654
PISA						
Outcomes						
Reading Scores	437.696	75.112	52,991	415.715	79.917	46,210
Mathematics Scores	410.145	70.114	52,991	425.791	75.005	46,211
Parental Investment Index	5.804	2.648	50,857	5.859	2.627	44,439
Attended Preschool (=1)	0.914	0.280	35,755	0.895	0.307	31,456

	Girls			**Boys**		
	Mean	S.D.	N	Mean	S.D.	N
Controls						
BaseRate (Diarrhea)	5.224	2.830	52,991	5.665	3.077	46,211
BaseRate (Respiratory)	3.374	2.031	52,991	3.857	2.292	46,211
BaseRate (Vaccine Preventable Disease)	0.077	0.101	52,991	0.095	0.107	46,211
Ln(GDP) in Birth State	6.536	0.548	52,991	6.554	0.556	46,211
Ln(Rainfall) in Birth State	9.384	0.408	52,991	9.387	0.416	46,211
Mother highest completed schooling ISCED level 1	0.188	0.391	51,728	0.179	0.383	44,995
Mother highest completed schooling ISCED level 2	0.283	0.450	51,728	0.295	0.456	44,995
Mother highest completed schooling ISCED level 3bc	0.161	0.367	51,728	0.158	0.365	44,995
Mother highest completed schooling ISCED level 3a	0.201	0.401	51,728	0.211	0.408	44,995
Mother highest qualification ISCED level 5b	0.154	0.361	47,161	0.154	0.361	40,907
Mother highest qualification ISCED level 5a6	0.152	0.359	47,937	0.176	0.381	41,951
Father highest completed schooling ISCED level 1	0.182	0.386	49,489	0.178	0.383	43,574
Father highest completed schooling ISCED level 2	0.271	0.444	49,489	0.277	0.448	43,574
Father highest completed schooling ISCED level 3bc	0.168	0.374	49,489	0.169	0.375	43,574
Father highest completed schooling ISCED level 3a	0.211	0.408	49,489	0.217	0.412	43,574
Father highest qualification ISCED level 5b"	0.131	0.337	45,522	0.142	0.350	39,978
Father highest qualification ISCED level 5a6"	0.146	0.353	46,559	0.164	0.370	41,023
% Girls in School	21.296	25.929	48,843	21.104	24.941	42,572
School Size	886.096	866.341	48,767	853.126	810.299	42,491
School Quality	-0.668	1.125	51,670	-0.707	1.124	45,050

Note: MxFLS 2002, 1987-1993 birth cohorts. See Notes to Table 3. Small birth size= 1 if the individual was reported by the mother as being (subjectively) small at birth. BaseRates are gender specific child (age 0-4) mortality rates per 1,000 births, averaged over the pre-intervention years, 1988-1990. In MxFLS, household head's education takes the values 0, 1, 2 for no primary, has primary and has secondary or higher education, respectively. In PISA 2003, 2006 and 2009, the sample consists of randomly selected 15 year olds tested in school and the internationally normalized scores have a mean of 400. The Parental Investment Index is the sum of binary indicators for these school-related household objects-desk, a quiet place to study, a personal computer, educational software, internet access, literature, poetry, textbooks, and a dictionary. For parent schooling, ISCED levels of education are: level 1=primary education, level 2=lower secondary, level 3a=upper secondary education giving direct access to theory based or research preparatory tertiary education, level 3bc=other upper secondary education. ISCED level 5b is a qualification from occupation-specific or practical first stage tertiary education programs and ISCED level 5a6 is a qualification from largely theoretical or research preparatory first or second stage tertiary education programs.

Table A2 -Testing for Differential Pre-Intervention Trends in Raven Scores Across High and Low Treatment States

	Girls	Boys
*Born in 1986**BaseRate	-0.680	-0.289
	(0.476)	(0.414)
*Born in 1987*BaseRate*	-0.811	0.00902
	(0.506)	(0.420)
*Born in 1988*BaseRate*	-0.763*	-0.264
	(0.426)	(0.343)
*Born in 1989*BaseRate*	-0.0705	-0.0874
	(0.539)	(0.481)
*Born in 1990*BaseRate*	-0.826	-0.0534
	(0.538)	(0.313)
N	2025	1908

Notes: Robust standard errors clustered at birth state level in parentheses. *** - p<0.01, ** - p<0.05, * - p<0.10. MxFLS data for the 1985-1990 (pre-intervention) birth cohorts. The dependent variable is the Raven score. The coefficients of interest are a vector of interactions between birth cohort and BaseRate (the pre-intervention birth state diarrheal mortality rate) and we control for birth cohort and birth state fixed effects. The birth cohort*BaseRate coefficients are jointly not significant. The one significant coefficient (for girls) is negative rather than positive so this pre-trend cannot explain our findings.

Table A3- Flexible Age of Exposure to the Clean Water Reform: Test Scores

	Girls				Boys			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Raven Scores								
*Exposure in year-2*BaseRate*	-0.0680	0.176	0.212	0.190	0.118	0.0971	0.167	0.636
	(0.234)	(0.218)	(0.189)	(0.263)	(0.188)	(0.339)	(0.338)	(0.426)
Effect size s.d	-0.011	0.029	0.035	0.031	0.021	0.017	0.030	0.114
*Exposure in infancy*BaseRate*	0.164	0.409	0.606*	0.518	-0.0624	-0.0868	-0.0976	0.451
	(0.252)	(0.299)	(0.342)	(0.314)	(0.190)	(0.334)	(0.310)	(0.464)
Effect size s.d	0.027	0.068	0.100	0.086	-0.011	-0.016	-0.018	0.081
*Exposure at birth*BaseRate*	0.430**	0.673***	0.957***	1.039***	-0.157	-0.189	-0.121	0.454
	(0.176)	(0.225)	(0.308)	(0.286)	(0.193)	(0.337)	(0.338)	(0.381)
Effect size s.d	0.071	0.111	0.159	0.172	-0.028	-0.034	-0.022	0.082
Reading Scores								
*Exposure in year-2*BaseRate*	2.519*	1.642	2.220		2.039	0.469	0.236	
	(1.434)	(1.646)	(1.626)		(1.729)	(1.795)	(1.904)	
Effect size s.d	0.098	0.070	0.089		0.167	0.112	0.107	
*Exposure in infancy*BaseRate*	2.368*	1.500	2.060		2.227	0.666	0.438	
	(1.239)	(1.381)	(1.353)		(1.534)	(1.758)	(1.894)	
Effect size s.d	0.092	0.065	0.083		0.092	0.037	0.032	
*Exposure at birth*BaseRate*	4.085***	3.143**	2.893**		1.971	0.375	-0.0670	
	(1.232)	(1.364)	(1.432)		(1.466)	(1.533)	(1.517)	
Effect size s.d	0.166	0.134	0.124		0.106	0.050	0.033	
Mathematics Scores								
*Exposure in year-2*BaseRate*	2.656**	1.770	2.103*		1.883	0.570	0.735	
	(1.171)	(1.210)	(1.220)		(1.469)	(1.392)	(1.387)	
Effect size s.d	0.100	0.074	0.087		0.062	0.020	0.035	
*Exposure in infancy*BaseRate*	2.857***	1.981*	2.293**		2.333*	1.026	1.174	
	(0.982)	(1.047)	(1.036)		(1.254)	(1.362)	(1.419)	
Effect size s.d	0.087	0.061	0.074		0.073	0.031	0.043	
*Exposure at birth*BaseRate*	4.508***	3.650***	3.239***		2.650**	1.310	0.700	

	(1.082)	(1.134)	(1.189)		(1.321)	(1.347)	(1.376)	
Effect size s.d	0.167	0.145	0.129		0.100	0.057	0.032	
Controls								
Birth State FE, Birth Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Household and School Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State*Year Controls	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Birth Region X Birth Year FE	No	No	Yes	Yes	No	No	Yes	Yes
Birth State Quadratic Trends	No	No	No	Yes	No	No	No	Yes

Notes: Robust standard errors clustered at birth state level in parentheses. *** - p<0.01, ** - p<0.05, * - p<0.10. See Notes to Table 3. Exposure at birth = 1 if the individual was exposed to the clean water program from the first day of life (April 1991 birth cohorts and thereafter), Exposure in infancy = 1 if the individual was exposed between the first and 12th month of life (April 1990 to March 1991 birth cohorts), Exposure in the year-2 of life = 1 if the individual was first exposed to the reform during the 13th-24th month of life (April 1989 to March 1990 birth cohorts). The omitted group is children first exposed after the age of 2 years. We have verified that there are no clean water exposure improvements in test scores for children who are first exposed at ages older than 2 years.

Table A4: Impacts of Exposure in Infancy to the Water Reform by Household Socioeconomic Status

Girls	**Low SES**				**High SES**			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Post*BaseRate	0.482	0.759**	1.010**	1.354***	0.108	0.189	0.0434	0.248
	(0.297)	(0.345)	(0.385)	(0.444)	(0.623)	(0.775)	(0.786)	(0.543)
Effect size s.d	0.077	0.121	0.161	0.216	0.016	0.027	0.006	0.036
N	1,595	1,595	1,595	1,595	932	932	932	932

Boys	**Low SES**				**High SES**			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Post*BaseRate	-0.272	-0.216	-0.111	0.279	0.347	-0.266	-0.366	-0.0218
	(0.234)	(0.465)	(0.431)	(0.411)	(0.386)	(0.401)	(0.507)	(0.444)
Effect size s.d	-0.048	-0.038	-0.020	0.049	0.051	-0.039	-0.054	-0.003
N	1,488	1,488	1,488	1,488	947	947	947	947
Controls								
Birth State FE, Birth Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Household and School Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State*Year Controls	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Birth Region X Birth Year FE	No	No	Yes	Yes	No	No	Yes	Yes
Birth State Quadratic Trends	No	No	No	Yes	No	No	No	Yes

Notes: Robust standard errors corrected for clustering at the birth state level in parenthesis. *** - p<0.01, ** - p<0.05. Dependent variable is Raven test score, sample is MxFLS. Equations are those in Table 3 but now the sample is stratified by the household head having secondary or greater education ("High SES") or not ("Low SES").

Table A5 - Impact Estimates by Rural v Urban Location of Household - Raven Scores

	Girls		Boys	
	Urban	**Rural**	**Urban**	**Rural**
Post*BaseRate	0.837**	0.788**	-0.0169	0.260
	(0.383)	(0.343)	(0.420)	(0.271)
N	1,498	1,029	1,433	1002

Note: Robust standard errors clustered at birth state level in parentheses. *** - p<0.01, ** - p<0.05, * - p<0.10. This is the specification in column 4 of Table 3 for rural and urban households respectively. Oportunidades was, for our sample cohorts, only available in rural areas.

Table A6: Impact of *In Utero* Exposure to Clean Water Reform on Birth Size

	Girls	**Boys**
Post*BaseRate	-0.00572	-0.00544
	(0.00998)	(0.00965)
N	1,074	1,071

Note: MxFLS 2002. Dependent variable is 1 if the mother reported the child was small at birth. Specification as in Table 3, column 4. Post is now adjusted to reflect births that were exposed to the water reform in utero. Mothers were asked the birth size question for their last two births.

Table A7 - Impact of Clean Water Program Exposure on Parental Schooling Investments Allowing Differential Impacts Throughout Early Childhood

	Girls	Boys
Parental Investment Index		
*Exposure in year-2*pre_diarrhea*	0.0399	0.0485**
	(0.0280)	(0.0229)
Effect size s.d	0.0400	0.0486
*Exposure in infancy*pre_diarrhea*	0.0533**	0.0619*
	(0.0242)	(0.0333)
Effect size s.d	0.0534	0.0620
*Exposure in birth*pre_diarrhea*	0.0360	0.0131
	(0.0267)	(0.0299)
Effect size s.d	0.0361	0.0131
Time on Homework		
*Exposure in year-2*pre_diarrhea*	0.160	-0.230***
	(0.0963)	(0.0732)
Effect size s.d	0.098	-0.157
*Exposure in infancy*pre_diarrhea*	0.218**	-0.0994
	(0.0960)	(0.0769)
Effect size s.d	0.133	-0.068
*Exposure in birth*pre_diarrhea*	0.203**	-0.104
	(0.0947)	(0.0938)
Effect size s.d	0.124	-0.071
Attend School		
*Exposure in year-2*pre_diarrhea*	-0.00655	0.00633
	(0.00647)	(0.00505)
Effect size s.d	-0.101	0.110
*Exposure in infancy*pre_diarrhea*	0.000696	0.00575
	(0.00258)	(0.00414)
Effect size s.d	0.011	0.100
*Exposure in birth*pre_diarrhea*	0.00125	0.000195
	(0.00308)	(0.00369)
Effect size s.d	0.019	0.003
Controls- All		

Note: Coefficients on exposure*BaseRate, where exposure is defined in Table A3. The investment variables are defined in Table A1. Controls as in the last column in Table 3. Robust standard errors clustered at birth state level in parentheses. *** - p<0.01, ** - p<0.05, * - p<0.10. Effect sizes are defined as in Table 3. See Table 4 for a more parsimonious specification corresponding to the main specification for test scores in Table 3. In this more flexible but more demanding specification, home work for girls becomes more clearly significant while school attendance loses significance. While birth year exposure continues to increase equipment for girls alone, we also see some positive effects of exposure in the first two years of life for boys. This is not inconsistent with our "story" but suggests that equipment investments for boys who were first exposed to the water reform after birth but before age 2 do not have large impacts on their adolescent test scores.

Data Appendix

The Mexican Family Life Survey (MxFLS) includes Raven test scores in addition to household demographics, expenditures, educational attainment, health and anthropometry, labor force participation, and fertility. The Raven test is an 18 item test for those aged 5-13 and a 12 item battery for those 14 and over. We present results using the 2001-2002 wave. While employing the 2005 wave offers gains in precision of the regression estimates, for many children of our sample cohorts, the same Raven's test was repeated in the 2005 wave. Thus, performance on the Raven's test in the second wave may reflect learning how to take this specific version of the test rather than meaningful changes in cognitive processing. We nevertheless obtained results for the 2005 wave alone and for the pooled sample and the broad pattern of results reported in the paper holds (available on request).

The Program for International Student Assessment (PISA) surveys drew a representative sample of school-going 15 year olds by first randomly choosing schools and then randomly choosing students within schools. PISA provide five different estimates of the test score for each testing domain. These represent plausible values from a posterior probability distribution delineated by PISA. We follow the OECD recommendation on combining these scores, using the *pv* command in Stata for the regressions. Unlike in the MxFLS, in the PISA data we only know the state of testing for each respondent, not the state of birth. Since over 80% of 15 year olds in the Mexican census currently reside in their state of birth (Venkataramani, 2009), we do not expect this to introduce a large bias.

The *Enquesta Nacional de Salud y Nutricion* (ENSANUT) 2012 survey was conducted by the Mexican School of Public Health between October 2011 and May 2012 (http://ensanut.insp.mx). Anthropometric measurements were taken from one randomly selected adult below the age of 60 in each household. We used these data in a companion paper to produce the reported results for heights of men and schooling of women in young adulthood (Bhalotra and Venkataramani 2013).

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*).[47] We also ran all equations in this paper using the infant mortality rate instead of the under-5 mortality rate and the results were very similar: infant mortality accounts for about two-thirds of under-5 mortality and the break in trend in the under-5 rate tracks the break in trend in the infant rate. We report results with the under-5 rate because of likely measurement error in infant mortality, which may systematically correlate with poverty and pre-program disease rates.[48]

---

[47] 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

[48] Infant deaths, most of which occur in the first few months of life, often went under-reported and cause of death reporting was thought to be inaccurate for infants(Tome, et al 1997).

Data on state GDP were taken from Vincente German-Soto (2005), pre-intervention literacy and schooling attainment in the state are from the 1990 Mexican Census (Steven Ruggles, et al, 2010), and rainfall data come from the National Weather Service website (http://smn.cna.gob.mx). We used the 2010 census microdata, at which date the marginal cohort (born 1990-1991) was age 19-20, allowing us to study survival rates in education overall and by gender. We obtained municipality level coverage of Progresa (Opportudinas) from Tania Barham and normalized this on municipality level population obtained from the 2010 census files.