

Barde, Julia Alexa; Walkiewicz, Juliana

**Conference Paper**

## The Impact of Access to Piped Drinking Water on Human Capital Formation - Evidence from Brazilian Primary Schools

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2013: Wettbewerbspolitik und Regulierung in einer globalen Wirtschaftsordnung

**Provided in Cooperation with:**

Verein für Socialpolitik / German Economic Association

*Suggested Citation:* Barde, Julia Alexa; Walkiewicz, Juliana (2013) : The Impact of Access to Piped Drinking Water on Human Capital Formation - Evidence from Brazilian Primary Schools, Beiträge zur Jahrestagung des Vereins für Socialpolitik 2013: Wettbewerbspolitik und Regulierung in einer globalen Wirtschaftsordnung, ZBW - Deutsche Zentralbibliothek für Wirtschaftswissenschaften, Leibniz-Informationszentrum Wirtschaft, Kiel und Hamburg

This Version is available at:

<https://hdl.handle.net/10419/79808>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

# The Impact of Access to Piped Drinking Water on Human Capital Formation

Evidence from Brazilian Primary Schools

March 1, 2013

We analyze the impact of access to piped water on human capital formation as measured by test scores from standardized school exams at Brazilian primary schools. We find a positive and significant effect of around 11 percent of the standard deviation of mean test scores. We find that the effect of piped water on test scores increases with the level of education of the mother and that this complementarity is more pronounced for families with income below average income and vanishes for families with income above mean. This allows important policy recommendations: focus infrastructure expansion on low income areas, complement them with educational interventions for families with low educational background.

**JEL-Codes:** I15, I25, H41

**Keywords:** Health, piped water, cognitive development, human capital formation

# 1 Introduction

A small but considerably growing literature analyses the consequences of lacking access to clean water on various health outcomes, as well as on educational and economic outcomes. It is part of the larger research area on the impact of health improvements on educational and economic outcomes. Results from both literature strands complement to make the link between access to clean water in the early years of childhood and economic success in adulthood via better health during early development and later schooling periods. Studies such as Fewtrell et al. (2005), Günther and Fink (2010) and Jalan and Ravallion (2003) show that there is a negative effect of increased access to improved or piped access to water on water related diseases such as diarrhea and nematodes and malnutrition of children. Water related diseases and malnutrition have been shown to impact negatively on years of schooling, school enrollment, school attendance and literacy (Bleakley, 2007; Bobonis et al., 2006; Miguel and Kremer, 2004). Other studies show the positive impact of health interventions on measures of economic success as wages and productivity (Alderman and Behrman, 2006; Baird et al., 2012; Fogel, 1994; Thomas and Strauss, 1997). The present study contributes to this literature by investigating the effect of access to piped water on educational achievements of children in a transition country.

We analyse the impact of access to tap water on schooling achievements using data from the school evaluation programme SAEB in Brasil from 1999 to 2005<sup>1</sup>. SAEB provides representative results from standardized tests in mathematics and Portuguese from all over Brasil and complements them with rich information on the socio-economic background of the children's families. This data allows us to contribute to the above literature in two ways. First, we look at the relationship between access to tap water and schooling achievement which has not been done so far. This approach has an important advantage when compared to other studies which focus on the effect of eridication of water related diseases. Miguel and Kremer (2004) for example find considerable positive short run effects of health interventions on health of primary children in Kenya. They find that only half a year after the distribution of deworming drugs, treated children report less incidence of water related diseases and go to school more often. Interestingly though, they find no significant improvement of test scores of the treated children. That is, the quantitative measure of human capital formation, absenteeism, is affected, but the health intervention does not affect the qualitative measure of human capital formation, test scores. One

---

<sup>1</sup>SAEB is called *Prova Brasil!* since 2007.

reason for this may be the short term focus of this study.<sup>2</sup> Gains from better health in terms of higher returns from school and learning may take longer to materialize. Health interventions evaluated only shortly after the experiment may not be suited to find these effects. Our variable of interest indicates whether a child has access to tap water at home at the time of the school exam. This variable can capture two effects. First, it may say something about the current or recent health status of the child. Or, second, if one assumes the current type of access to water of a family to proxy the past type of access, as well, our variable may also capture long term effects of extended periods of water related diseases during childhood. Some studies have evaluated the effects of early childhood health shocks on educational outcomes<sup>3</sup>. The focus of these studies was however only loosely related to water related diseases and the outcome variables were again quantitative measures of educational attainment such as years of schooling or enrolment rates. Our study aims at providing new evidence for the link between access to piped water and human capital formation. We therefore chose a dependent variable which depicts qualitative information about human capital, and an explaining variable which allows long term effects of reduced incidence of water related diseases to show up.<sup>4</sup>

Our second contribution lies within the analysis of the heterogeneity of the effects of access to piped water. Only very few contributions have focused on this issue so far and to the best of our knowledge none has studied the heterogeneity of effects on educational attainment. Heterogenous effects of income and education of mothers have also been found for the impact of tap water access on health outcomes of young children (Gamper-Rabindran et al., 2010; Jalan and Ravallion, 2003). As piped water needs a large and expensive infrastructure it is important to analyze under which conditions the returns to such investments are largest. We find that children from poorer families benefit more from access to tap water than families at mean or above mean income. However, this effect is conditional on at least some years of schooling of the mother.

The fact that we can contribute into these two directions to the literature comes at the expense of a methodologically absolutely safe identification strategy. As many others, we

---

<sup>2</sup>A complementary study to this paper by Baird et al. (2012) finds significant and large effects in the long run when comparing treatment and control groups 12 years after the initial intervention. However, the focus is on labour market participation as the by then adult persons left school already.

<sup>3</sup>See section 3 a review with references.

<sup>4</sup>The drawback of our variable of interest (as compared to health intervention treatment dummies) is that we would need additional information about the actual health of the child at the date of the test and in the past to identify the health channel between access to water and school achievements that the literature suggest. Section two will further elaborate on this issue explaining why access to tap water is a valid proxy for health.

don't dispose of readily available instruments for access to water, and our data is neither experimental nor does it allow for regression discontinuity design. A large part of our paper therefore focuses on how to control for the socio-economic background of the children which probably affects the probability that they have access to water as well as their performance at school positively. We construct a proxy for permanent family income in order to control for economic background of the children and include several other variables, such as education of the parents, whether the child has to work out of school, or availability of electricity at home. Our results are robust to the inclusion of these variables and we can identify significant separate effect. We further allow for more functional flexibility and run several robustness checks for different subsamples. The fact that our variable of interest only marginally changes in size but never in significance through these tests and that also the analysis of the effect heterogeneity delivers a plausible story, eases our concerns about endogeneity to some extent. Although one has to keep in mind that the effects that we find might still be upward biased, we think that our study contributes important insights to the literature and allows for formulation of careful policy implications and shows important directions for further research on the link between access to piped water and the formation of human capital.

Our paper is organized as follows. Section two further elaborates on the link between access to clean water and educational achievement. Section three explains the data we use and section four presents the results. Section five presents the results on effect heterogeneity and section six concludes.

## 2 Water, Health & Educational Achievement

Equation 1 gives a simple production function approach to educational achievement.<sup>5</sup> The test score  $TS$  of child  $i$  in an exam administered at school is determined by its health  $H$ , parental inputs into education  $P$ , school inputs  $S$ , years of schooling  $YS$  and its innate ability  $\mu$ .

$$TS_i = f(H_i, P_i, S_i, YS_i, \mu_i) \quad (1)$$

We choose test scores as a measure of educational achievement. This choice puts more weight on the quality aspect of educational attainment than on the quantitative aspect. Quantitative measures could be completed years of schooling or absenteeism in days. Equation 1 captures the direct impact of health on educational achievement which can work either through an increase in the time spent at school and/or learning, or through an im-

---

<sup>5</sup>See Glewwe and Miguel (2008) for a more detailed exposition of this function.

provement of the productivity in learning.<sup>6</sup> We estimate a specification based on equation 1 in section 3 and 4 of this paper as we think that access to tap water affects educational achievement via health. Water related diseases such as intestinal worms or diarrhea are very common in developing countries and especially harmful to young children below five who also attract them more often than older children or adults. Glewwe and Miguel (2008) calculate that 17% of the healthy years lost by children aged zero to four because of diseases are lost because of diarrhea (12.6%) or other nutritional distortions (4.4%).<sup>7</sup> Prüss-Üstün et al. (2008) estimate that 50% of all malnutrition of children is due to water related diseases. Whereas it is clear that access to improved water and sanitation, or even better piped water and piped sanitation, very effectively reduces the incidence and duration of water related diseases (Günther and Fink, 2010; Jalan and Ravallion, 2003; Kremer et al., 2011), there is no direct evidence that water related diseases and their consequences affect human capital formation negatively. However, taken together, the literature on the consequences of water related diseases for health and the literature on the consequences of reduced height for educational attainment and later economic outcomes, are very suggestive. Apart from the acute symptoms, the permanent consequence of chronic undernutrition due to frequent diarrhea and anemia is especially stunting, a negative deviation from the average height for age. Numerous epidemiological studies document permanent effects of diarrhea and malnutrition on physical growth of young children (Checkley et al., 2008; Dillingham and Guerrant, 2004; Guerrant et al., 2002; Moore et al., 2001) and also evidence from randomized experiments in developing countries by economists starts to confirm this relationship (Bobonis et al., 2006).<sup>8</sup> Further, there is abundant literature showing causal effects of height for age below five on educational attainment and labor market outcomes.<sup>9</sup> Case and Paxson (2008) show that in particular height for age is a proxy of cognitive abilities of young children and that it is for this very reason that it impacts on so many outcome

---

<sup>6</sup>Equation 1 understates the role of the parents especially in the early years of childhood who are responsible of providing “health” to their children. Starting with a standard utility function of the parents, one can derive a conditional demand function for educational achievements which contains their demand for health inputs and also their responses to changes in health of their child. The indirect effects cannot be captured by the production function approach above. See Glewwe and Miguel (2008) for a more comprehensive exposition. We will come back to this issue later.

<sup>7</sup>For children aged five to 14, 8.4% of the total burden can be explained by water related diseases. Death is not included into the calculation of healthy years lost.

<sup>8</sup>Bobonis et al. (2006) find that a reduction of helminths infections leads to weight gains of young children in India.

<sup>9</sup>See the reviews by Almond and Currie (2010) and Glewwe and Miguel (2008) and the references therein.

variables.<sup>10</sup> We therefore feel reasonably safe to state that water related diseases impact on cognitive development of children through their impact on health. The present study aims at measuring this effect by testing whether children who have access to tap water systematically have better results in achievement test.

There is some recent evidence from developing countries which focuses on the impact of randomized health interventions on educational outcomes and contributes first insights on the role of water related diseases.<sup>11</sup> Bobonis et al. (2006) and Miguel and Kremer (2004) find that a reduction in the incidence of intestinal worm infections and/or the daily administration of iron and vitamine A supplements lead to a sharp increase in school attendance of the treated children and even of their non-treated fellows in the short run. Miguel and Kremer (2004) additionally show that this increase is accompanied by a reduction in reported diarrhea incidence and Bobonis et al. (2006) find gains in weight for treated children. Both studies use a quantitative schooling measure as dependent variable. If one assumes that more time spent in school increases the amount of knowledge and capacities of children or young adults, this evidence shows a positive relationship between nutrition and/or water related diseases and academic achievement. Miguel and Kremer (2004) also analyze whether deworming affects qualitative measure of educational achievement. They evaluate exams that took place about 12 months after the treatment started. They find no significant effects. The authors suggest based on their findings that the increase of school attendance that they find is not sufficient to improve academic performance because already 12 months after the treatment the helminths infection burden is back to 80% of its initial size and attendance probably decreases again. Miguel and Kremer (2004) suggest that it needs a permanent increase in attendance over a long time to expect an impact of improved health on test scores. The literature reviewed above further suggests, that the impact of health on achievement does not only work through attendance but also through the development of cognitive skills in early childhood. We hope to find these effects by choosing access to tap water as a proxy variable for current and, most importantly, past health status. To the best of our knowledge, there is no further evidence from developing countries addressing the link between qualitative improvements in school achievement and water related diseases.

---

<sup>10</sup>Case and Paxson (2008) reviews the epidemiological literature with respect to question why physical growth and the development of cognitive abilities are influenced by the same external factors. Their does not seem to be an answer to this question yet.

<sup>11</sup>There is other literature suggesting an effect of nutrition and health on educational attainment (e.g. Alderman et al. (2006, 2009)), however it does not directly allow to learn anything about the impact of water related diseases. See Glewwe and Miguel (2008) for an extensive review of empirical literature of older literature on health and education.

## 3 Data & Research Design

### 3.1 Data

We use data from the *Sistema Nacional de Avaliação da Educação Básica* (SAEB), the national education evaluation programme implemented every two years by the Brazilian Ministry of Education. SAEB contains individual results from nationwide, standardized tests in mathematics and Portuguese in the fourth and eighth grade of the Brazilian *ensino fundamental* and in the third grade of the *ensino medio*. We focus on the test results from fourth grade which is equivalent to fourth grade of European or US primary school: Children are on average 10.8 years old. The sampling strategy of SAEB allows for representative results at the national, rural and urban level. Urban data is also representative at the level of the five official Brazilian regions and the 27 federal states. The rural sample is considerably smaller and therefore only representative at the regional level. SAEB is a rotating school panel. Overall, the sample of fourth graders of all four years contains 9200 schools and around 12 pupils per school and discipline<sup>12</sup>. The standardized tests administered by SAEB are designed such that the test scores of children are comparable between all rounds. The exam questions are primarily meant to test for cognitive capabilities of the children, i.e. all questions relate to a specific cognitive capability, such as e.g. applying a standard solution technique to a new and/or slightly different problem or draw a conclusion from a text (SAEB, 2006).<sup>13</sup>

In addition to the tests, children, teachers and directors fill out questionnaires about complementary information describing the childrens' home environment, learning experience at home and school, schooling inputs and teachers quality. The data set contains information ranging from variables measuring education of the parents, infrastructure availability at home, or childrens' interests out of school, or teacher's education, the availability of computers at school or the condition of the schools' bathrooms. Our variable of interest, whether a children has access to tap water, was asked in the survey waves of 1999, 2001, 2003 and 2005. 1999 is also the first year of the survey available publicly, in 2007 the

---

<sup>12</sup>In total, there are around 24 pupils per class in the sample. The test is randomized within schools, i.e. the class which takes the test is chosen randomly if there is more than 1 class per level. If there are three or more than three classes, 2 classes are chosen. Within one class, half of the pupils is randomly selected to take the test in mathematics. The other half takes the test in Portuguese. See SAEB (2008) for further details.

<sup>13</sup>Eventhough they aim at measuring cognitive capabilities, also the non-cognitive skills of the children, such as their ability to concentrate for a given time or their patience, influence their test results. Also external events, such as a lack of sleep at the night before the exam or also the teacher-pupil relation can influence the performance of the child. Cf. Cunha and Heckman (2007)



design of the questionnaire was changed and the question dropped.<sup>14</sup> The question focuses explicitly on *piped* water: In 1999 and 2001 the question was “Is there piped water where you live?” (“*Onde você mora existe água encanada?*”) and in 2003 and 2005 “Is there tap water where you live?” (“*Onde você mora chega água pela torneira?*”).<sup>15</sup> The questions only asks about the “*place where you live*”, i.e. it is not clear whether children state with their answer that their is piped water from a tap within their house or flat, or whether their is piped water in the courtyard, building or plot (but not within their house or flat). As highlighted in section 2, we expect a positive impact of access to *piped* water on schooling achievement.

Overall, 89.7% of the children in the sample have access to tap water. However, there are large intertemporal and regional differences. table 1 shows the traditional differences between the rather rich regions South-East and South, and the rather poor North and North-East in Brazil. In the richer regions almost all children answer that they have tap water at home in 2005, in the poorer regions only about 85% of the children say the same in the same year. As expected, access increases over time, but unevenly across the regions and over time. Especially, the 2001 wave shows some peculiarities. First, in almost all five regions, access is reported to be lower in 2001 than in 1999. Second, the jump from 2001 to 2003 is considerably larger in South-East, South and Centre. There have been some changes in the sampling strategy between the 2001 and the 2003 wave, however none of these changes could in our opinion explain the pattern in table 1.<sup>16</sup> This raises concerns of comparability of the four waves. We will run several robustness checks to address these concerns.

Figures 1, 2 and 3 plot the cumulative density of access shares per school for different samples. Figure 1 shows that around 55% of all schools have access shares below 100%. We take from this figure that variation is well distributed accross schools. Figure 2 shows the split between the rural and urban schools. Figure 3 shows the distribution for the four

---

<sup>14</sup>Since 2007, SAEB is called *Prova Brasil!*

<sup>15</sup>See e.g. SAEB (2004) or SAEB (2008) for the documentation of the surveys.

<sup>16</sup>The biggest change between 2001 and 2003 is a considerable increase in the areas taking into account for rural sampling. In 1999 and 2001, rural schools were tested and interviewed only in the federal states of Minas Gerais and Mato Grosso do Sul and in the North-East region. In 2003, however, schools in rural areas of all states were included into the sample. It is a priori unclear whether the inclusion of these areas increased the average availability of public water infrastructure. When compared to the urban and rural areas included already in 1999 and 2001, the newly added regions are not known to be all better (or worse) served with infrastructure. The North e.g. is known to be the least equipped with public infrastructure, the states of Rio de Janeiro or Sao Paulo are the most developed federal states in Brazil and have been so for a long time. Thus, the above pattern might be explained by this change in sampling but we don't know for sure.

Table 1: **Regional and intertemporal differences in water access**

	North	North-East	South-East	South	Centre	Total
<b>1999</b>	0.777	0.829	0.892	0.917	0.879	<b>0.867</b>
<b>2001</b>	0.805	0.805	0.879	0.909	0.886	<b>0.856</b>
<b>2003</b>	0.836	0.836	0.975	0.983	0.947	<b>0.920</b>
<b>2005</b>	0.851	0.877	0.987	0.988	0.953	<b>0.941</b>
<b>Total</b>	<b>0.820</b>	<b>0.836</b>	<b>0.933</b>	<b>0.951</b>	<b>0.916</b>	

Note: The table shows average access to tap water at home.

Source: own calculations with SAEB 1999-2005.

waves. Again, the 2001 wave is different from the three others. Especially, the density above the access share of 60% seems to be considerably higher than in the three other waves.<sup>17</sup> This re-confirms our concerns from above.

Figure 1: Access share per school, total sample

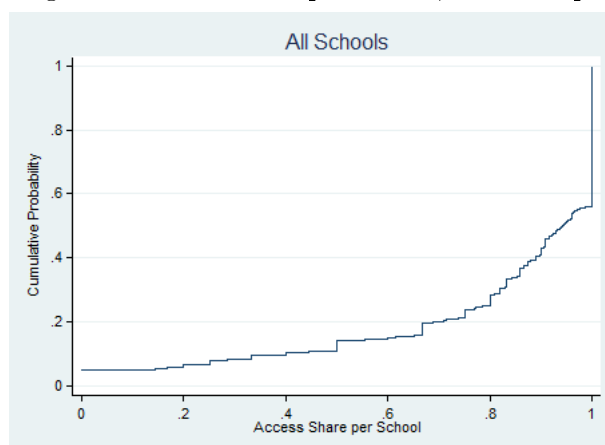


Table 2 shows descriptive statistics about the socio-economic background of the children and their families. It distinguishes between children at rural and urban schools. The first column of the rural and urban sections gives the sample means for all children at these schools. The second and third row in each section separates children with access to tap water from children without tap water.

All indicators of socio-economic background and child characteristics are significantly different between the rural and urban sample. Fourth graders are older and have worse test results in rural areas. Urban parents are on average better qualified, they have better in-

<sup>17</sup>Again, this might be due to the sampling change for the waves of 2003 and 2005. As more rural schools were added to the sample and those probably have less access than urban schools already in the sample, at least the shift in the right hand part of the graph might be explained. The distribution of schools gets flatter in 2003 and 2005 across all access rates.

Figure 2: Access share per school, rural vs. urban

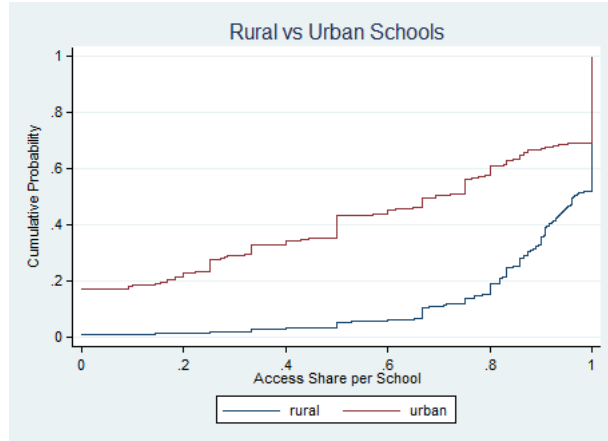
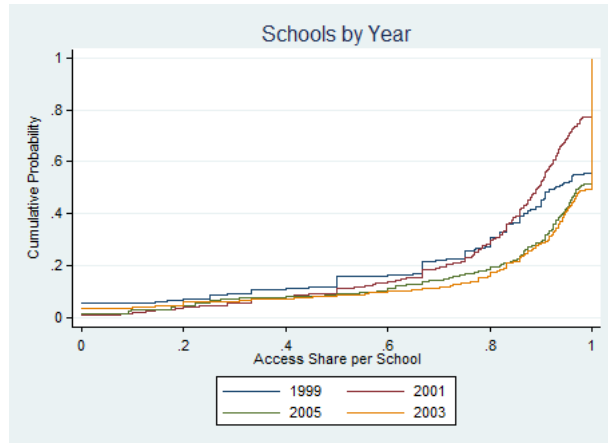


Figure 3: Access share per school, per year



frastructure and less of their children work before or after school. In rural areas especially European and indigenous families have tap water access at home. This is also true in urban areas, however differences are less pronounced here.

Splitting the sample into children with access to tap water and children without tap water at home reveals that these two groups also have quite different family backgrounds in rural as well as in urban areas. Children having access to tap water are younger in both areas and test scores of children with tap water are better. Electricity is better available than tap water: Even 89% of the children not having tap water at home, report that there is electricity at home. In rural areas, electricity coverage is lower and the difference between both samples is larger. Mothers in families with piped access to drinking water are better educated than mothers of families without access to tap water at home. This is also true for fathers. 25% (44%) of the children without water access in urban (rural) schools work before or after school. These shares are considerably lower for children that have access

Table 2: Descriptive statistics, 4th grade, 1999-2005

	URBAN			RURAL			min	max	mean
	all	no DW	DW	all	no DW	DW			
European			0.93			0.72			
Mulatto			0.92			0.67			
Black			0.90			0.57			
Asian			0.90			0.63			
Indigenous			0.93			0.73			
Age	10.73	11.13	10.70	11.44	11.81	11.26	8.0	15.0	10.8
Test score	186.63	170.54	188.01	160.35	154.68	163.12	66.7	373.4	184.0
Electricity	0.97	0.89	0.98	0.85	0.71	0.92			
Mum_noeduc	0.05	0.09	0.05	0.13	0.20	0.11			
Mum_primary	0.26	0.31	0.26	0.44	0.42	0.44			
Mum_secondary1	0.19	0.19	0.19	0.12	0.10	0.13			
Mum_secondary2	0.11	0.08	0.12	0.05	0.03	0.06			
Mum_university	0.12	0.07	0.13	0.03	0.02	0.03			
Dad_noeduc	0.06	0.10	0.05	0.16	0.22	0.14			
Dad_primary	0.20	0.25	0.20	0.32	0.30	0.34			
Dad_sec1	0.15	0.15	0.15	0.10	0.07	0.11			
Dad_sec2	0.10	0.06	0.10	0.04	0.03	0.04			
Dad_university	0.12	0.07	0.13	0.02	0.01	0.03			
Child works	0.15	0.25	0.14	0.38	0.44	0.35			
Books at home	0.82	0.77	0.83	0.74	0.67	0.78			
Income	0.21	-0.59	0.28	-1.75	-2.53	-1.37	-4.86	7.13	0.01
observations	115,802	8,490	107,312	9,452	2,861	6,591			

to piped water. When asked whether their family disposes of books at home (at least one book), 77% percent of the urban children without water access say yes as compared to 83% of the children with water access. In rural schools the difference is again larger.

The last variable in table 2 proxies for long term wealth of the families. As we don't observe true income of the children's families, we construct a long term wealth proxy for all families using principal component analysis. The variables we include into the construction of the index are household size, persons per room, whether the family disposes of domestic help, whether there is a bathroom, the number of cars, and the existence of the following assets: TV, radio, video, PC, fridge, freezer, vacuum cleaner, cloth washer. This procedure has two drawbacks. First, all of these questions but the one for a bathroom at home depict wealth at a relatively high level. As questions such as e.g. for the quality of walls and floors or sanitation type are not included in the survey we cannot capture real poverty as defined e.g. by the DHS surveys. Poverty would be more than not having a domestic help or a fridge at home. However, we are confident that some variables included in SAEB help to reduce this bias. We dispose of information about electricity availability at home and about the existence of a bathroom within the house/flat. In 1999, children also state whether there is a paved road leading to their home. Obviously, also our variable of interest, the availability of tap water of home is a comparable proxy of the long term economic

situation of households. Housing experts even consider variables depicting the immediate environment of the household (existence of public illumination, type and number of paved roads, sanitation type) as better suited to depict poverty in Brasil than questions about the type of walls and floors (IBGE, 2008). Our aim is to identify the effect of tap water access on school achievement and especially, to isolate it from income effects. The (socio)-economic background of the families which drives the probability to have piped water at home as well as schooling achievement is a threat to identification. We therefore have to be careful about the construction of the income index and the inclusion of additional control variables. We decide to include the bathroom indicator into the principal component analysis and control for electricity availability separately. As the indicator variable of paved roads is just available for 1999, we will run additional robustness checks with this variable on the reduced smaller sample. The problem about the bathroom variable is that we cannot be sure what exactly is measured by this variable. The SAEB methodology gives no further indication what exactly a bathroom should be like (should there be a toilet? if so, should it be connected to waste water collection? should there be a shower or a bath tub?) and also no further information is given to the children. Some studies show that they are strong complementary effects of appropriate private sanitation facilities and tap water availability. We will include this variable later as a separate regressor (and then exclude it from the wealth indicator)<sup>18</sup>, but for now we prefer to include it into the principal component analysis to avoid (unclear) confounding effects to our main variable of interest.<sup>19</sup> The second drawback of using a principal component analysis is that we almost only dispose of dummy variables. Principal component analysis has been shown to be better suited for continuous variables. However, it has been applied to the problem at hand very often, as well (see e.g. McKenzie (2005)), and there is no competing procedure we could choose.<sup>20</sup> From table 2 we see that our wealth proxy depicts considerable variation in our sample. The mean income in rural and urban areas is statistically different, and also the average income of families with tap water at home in both areas is larger than the income of families without tap water.

---

<sup>18</sup>Not contained in this version yet.

<sup>19</sup>Unclear effects could stem from the uncertainty about what we measure with this variable and the high correlation between this indicator and our variable of interest.

<sup>20</sup>The Brazilian Ministry of Education proposes a weighting and pricing scheme to construct income figures (in R\$) from SAEB data. We will use the income index resulting from this procedure as robustness checks in a later version of the paper.

### 3.2 Research Design

Our dependent variable is the mathematics test score of child  $i$  in school  $s$ .<sup>21</sup> We standardize the test scores  $T_{is}$  to mean zero and standard deviation one to allow for comparison of the coefficients. SAEB does not allow to follow individual pupils over time. We therefore look at repeated cross sections of fourth graders controlling for school fixed effects,  $\mu_s$ , and year dummies,  $t_i$ . Further, we include control variables  $X_i$  depicting the economic background of the child. We discussed these variables above. We cluster the error terms at the school level to allow for correlation of pupil from the same school. Equation (1) summarizes our approach.

$$T_{is} = \alpha + \beta_1 Water_i + \beta_2 X_i + \mu_s + t_i + \epsilon_i \quad (2)$$

The main concern of our regression model is that the error term contains some unobserved component which drives the probability of having tap water at home and the test performance in the same time. This would invalidate our coefficient estimate. Whether a family has tap water at home depends on two factors. First, the municipality where the family lives has to dispose of a functioning distribution network for fresh water which connects all neighbourhoods to the utility. Whether this is the case is a question of the wealth/development of the municipality and also of the willingness of political leaders to connect poor or even illegal settlements to the grid. Second, if water access is provided in principle by the municipality, the family has to decide whether to uptake the connection and bear the costs for water consumption. If water access is not available in the current place of living, the family has to decide about moving to a connected neighbourhood. Both decision are a function of the family's income. The higher income, the higher the probability to live in a connected neighbourhood and the higher the ability to afford uptake and consumption fees. The first factor, wealth or development status of the municipality where the family lives, might also impact on school quality through better institutional quality. As school quality influences school achievements of children, this is a potential source of upward bias of our coefficient estimates. We address this concern by including school fixed effects. As long as there is no variation in school quality over time which is correlated through institutional changes to network expansion, school fixed effects take care of the cross sectional differences. We think that the development stage of a municipality and the quality of its institutions which lead to improvement in public good provision (education and water supply) is changing only relatively slowly. In our baseline specification we look

---

<sup>21</sup>This paper contains only the results for mathematics. Test scores from the Portuguese exams will be included in a later version. Preliminary results show that the results are not different and therefore will serve as robustness check.

at a time span of eight years. Our identifying assumption might be only weakly satisfied for this long time period. In the robustness checks to our main specification we will therefore reduce this time span to six and four years.<sup>22</sup> The second parameter which threatens identification in the above specification is the economic background of the parents. We have discussed this issue in the data section above, and will include several proxies for the socio-economic background of the families to reduce this source of endogeneity. The above descriptives show that our control variables capture considerable variation between the different samples.

## 4 Results

Table 3 presents our baseline specification. It contains a number of à priori exogenous variables which we include in all of the following specifications.<sup>23</sup> The first four variables account for the ethnic background of the child. Considering oneself as from European descent (“white”) is the omitted category. Being black or from the indigenous community is significantly negatively correlated with schooling results as compared to the reference category. Next, we control for age and sex of the children. As we focus on the results of mathematical tests, the highly significant negative coefficient for female pupils is not surprising. Age is correlated significantly negative with test scores. This result is counter-intuitive at first sight. Especially during the first years at school, age has been reported to impact positively on school achievements. In a variation of this first baseline regression (not shown) we included dummies for each age category. Whereas children aged 7-12 perform significantly better than the omitted category (age 6), children older than 12 perform significantly worse than their younger class mates. In the specification underlying table 3 this negative effect seems to outweigh the positive one.

After controlling for time dummies, we add our variable of interest to this first specification. Having access to tap water at home turns out to be positively and significantly related to test scores and (if taken to be causal) explains 14% of the standard deviation in test scores. This effect remains robust to the inclusion of electricity. Electricity is also positively and significantly related to test scores, the coefficient is even twice as high as the water coefficient. Also adding the income proxy to the regression does not alter our

---

<sup>22</sup>In a later version of the paper we will also include variables that have been shown to measure school quality, such as teacher-pupil-ratios or availability of infrastructure and pedagogical material. However, already the descriptive statistics show that these variable change only marginally over time.

<sup>23</sup>Unless otherwise mentioned, the coefficients of these variables don't change in magnitude or significance. We don't show the results for these variables in the following in order to keep tables legible.

Table 3: **Baseline Results**

<b>Dependent Variable: Test Score Math, Grade 4</b>				
mulatto	-0.00923 (0.0141)	-0.00691 (0.0142)	-0.00741 (0.0141)	-0.00842 (0.0141)
black	-0.261** (0.0194)	-0.258** (0.0197)	-0.255** (0.0197)	-0.254** (0.0217)
asian	-0.0586 (0.0301)	-0.0502 (0.0291)	-0.0455 (0.0283)	-0.0470 (0.0322)
indigenous	-0.0604* (0.0260)	-0.0614* (0.0262)	-0.0622* (0.0261)	-0.0544 (0.0291)
female	-0.0724** (0.0106)	-0.0719** (0.0107)	-0.0697** (0.0107)	-0.0698** (0.0114)
age	-0.0913** (0.00390)	-0.0906** (0.00397)	-0.0895** (0.00398)	-0.0845** (0.00437)
2001	-0.0738** (0.0270)	-0.0636* (0.0271)	-0.0615* (0.0274)	-0.00400 (0.0297)
2003	-0.0919** (0.0325)	-0.103** (0.0327)	-0.103** (0.0330)	-0.131** (0.0360)
2005	0.0541 (0.0340)	0.0447 (0.0340)	0.0466 (0.0342)	0.0584 (0.0358)
tap water		0.149** (0.0169)	0.124** (0.0174)	0.130** (0.0205)
electricity			0.266** (0.0264)	0.240** (0.0302)
income				0.0359** (0.00492)
Constant	1.127** (0.0512)	0.993** (0.0533)	0.751** (0.0586)	0.769** (0.0657)
Observations	154,756	152,391	150,709	124,167
Adjusted R-squared	0.343	0.345	0.347	0.350

Note: Significance levels: \* $<0.05$ , \*\* $<0.01$ . All specifications contain school fixed effects and are clustered at the school level.



results. The coefficient of the income proxy is relatively small. This might be due to the fact that (almost) all variables summarized with the first principle component are proxies at a relatively high level (see discussion above). The fact that we can distinguish water and electricity effects separately from other income effects, gives us a first reason to believe that endogeneity is at least reduced by our specification choice. This result will be reinforced by the following specifications in table 4 that further control for strong indicators about the socio-economic background of families by adding more measures of ability and economic success of the parents. First, we add the highest educational level the mother could attain. Every educational level of the mother is as expected positively related to childrens' test scores when compared to non educated mothers. The coefficient increases with the time span mothers spent at school. The coefficient for mothers who graduated from highschool and attended university is smaller than mothers who graduated at most from highschool. Fathers' education turns out to be mostly insignificant.<sup>24</sup> The next variable we include, "child works", indicates whether a child works at home or outside the parental house before or after school and therefore have to contribute to either the household's income or to substitute for a help that the household cannot afford. This is correlated negatively and highly significant to test scores. The last variable in table 4 indicates whether the child thinks that they are more than 20 books at home. The correlation is as expected positive and significant. Even though we have included many strong indicators of socio-economic backgrounds of the children, the coefficient of tap water at home remains highly significant throughout all specifications and does almost not alter in magnitude. It stabilizes around 11% of the standard deviation.

We now turn to several robustness checks to the last specification of table 4. The first column of table 5 repeats its results. Column two and three show the results of the same specification run on the rural and the urban school sample separately. As shown in the descriptive statistics, children's characteristics, test results and families are very different in rural and urban areas. Therefore we run the same specification as above now on two separate samples allowing for heterogeneous effects of all variables. Column 2 and 3 in table 5 show the results. Whereas the urban sample shows the same results as the full sample (and some coefficients even increase), the rural sample only shows significant results for

---

<sup>24</sup>The question for parents' education contains the category "don't know" which shows to be significantly and positively correlated to test scores for mothers' and fathers' education. One possible explanation could be that children with higher school achievements check the "don't know" category more often than children with lower test scores who leave this question unanswered because the former understand better what the "don't know" category is supposed to mean whereas the latter don't know how to react and leave it open.

Table 4: **Base + Home Environment**

<b>Dependent Variable: Test Score Math, Grade 4</b>					
tap water	0.122** (0.0207)	0.122** (0.0207)	0.114** (0.0207)	0.114** (0.0208)	0.113** (0.0208)
electricity	0.235** (0.0307)	0.238** (0.0313)	0.218** (0.0316)	0.217** (0.0316)	0.218** (0.0316)
income	0.0257** (0.00494)	0.0258** (0.00494)	0.0274** (0.00499)	0.0278** (0.00499)	0.0256** (0.00497)
Mum_primary	0.106** (0.0287)	0.0959** (0.0302)	0.0887** (0.0301)	0.0908** (0.0303)	0.0918** (0.0304)
Mum_secondary1	0.177** (0.0321)	0.152** (0.0329)	0.144** (0.0324)	0.147** (0.0326)	0.146** (0.0327)
Mum_secondary2	0.340** (0.0328)	0.297** (0.0348)	0.285** (0.0343)	0.289** (0.0345)	0.287** (0.0346)
Mum_university	0.221** (0.0347)	0.190** (0.0361)	0.182** (0.0361)	0.184** (0.0364)	0.177** (0.0364)
Mum_don't know	0.112** (0.0293)	0.0788* (0.0306)	0.0701* (0.0303)	0.0743* (0.0306)	0.0745* (0.0307)
Dad_primary		0.0308 (0.0275)	0.0330 (0.0277)	0.0345 (0.0279)	0.0339 (0.0280)
Dad_secondary1		0.0471 (0.0302)	0.0508 (0.0304)	0.0533 (0.0305)	0.0521 (0.0305)
Dad_secondary2		0.123** (0.0347)	0.122** (0.0348)	0.123** (0.0349)	0.120** (0.0350)
Dad_university		0.0583 (0.0316)	0.0575 (0.0318)	0.0597 (0.0320)	0.0558 (0.0320)
Dad_don't know		0.0676* (0.0273)	0.0653* (0.0276)	0.0664* (0.0277)	0.0655* (0.0277)
child works			-0.190** (0.0169)	-0.190** (0.0170)	-0.191** (0.0169)
with parent(s)				0.0597** (0.0189)	0.0588** (0.0189)
books					0.0691** (0.0155)
Constant	0.558** (0.0728)	0.521** (0.0751)	0.536** (0.0746)	0.469** (0.0768)	0.469** (0.0766)
Observations	119,894	117,901	116,989	116,323	116,323
Adjusted R-squared	0.356	0.359	0.363	0.363	0.364

Note: Significance levels: \* < 0.05, \*\* < 0.01. All specifications contain school fixed effects and are clustered at the school level.

Table 5: **Robustness checks**

<b>Dependent Variable: Test Score Math, Grade 4</b>				
	<b>all</b>	<b>rural</b>	<b>urban</b>	<b>03/05</b>
tap water	0.113** (0.0208)	-0.0108 (0.0515)	0.140** (0.0224)	0.0696* (0.0290)
electricity	0.218** (0.0316)	0.0929 (0.0579)	0.280** (0.0355)	0.247** (0.0345)
income	0.0256** (0.00497)	0.00795 (0.0148)	0.0271** (0.00519)	0.0254** (0.00568)
Mum_primary	0.0918** (0.0304)	0.0957 (0.0613)	0.0955** (0.0344)	0.0919* (0.0378)
Mum_secondary1	0.146** (0.0327)	0.150 (0.0790)	0.150** (0.0363)	0.152** (0.0416)
Mum_secondary2	0.287** (0.0346)	0.323** (0.0704)	0.288** (0.0383)	0.299** (0.0427)
Mum_university	0.177** (0.0364)	0.172 (0.113)	0.181** (0.0394)	0.157** (0.0420)
Mum_don't know	0.0745* (0.0307)	-0.0135 (0.0638)	0.0854* (0.0344)	0.0644 (0.0370)
Dad_primary	0.0339 (0.0280)	0.0443 (0.0507)	0.0460 (0.0320)	0.00974 (0.0348)
Dad_secondary1	0.0521 (0.0305)	-0.133 (0.0824)	0.0799* (0.0332)	0.0324 (0.0380)
Dad_secondary2	0.120** (0.0350)	0.0537 (0.101)	0.140** (0.0382)	0.0877* (0.0433)
Dad_university	0.0558 (0.0320)	0.0625 (0.139)	0.0721* (0.0348)	0.0172 (0.0374)
Dad_don't know	0.0655* (0.0277)	0.00117 (0.0582)	0.0851** (0.0313)	0.0273 (0.0335)
child works	-0.191** (0.0169)	-0.148** (0.0409)	-0.195** (0.0186)	-0.196** (0.0198)
with parent(s)	0.0588** (0.0189)	0.0463 (0.0477)	0.0586** (0.0200)	0.0100 (0.0258)
books	0.0691** (0.0155)	-0.0460 (0.0801)	0.0745** (0.0157)	0.0625** (0.0177)
Constant	0.469** (0.0766)	0.229 (0.174)	0.422** (0.0841)	0.540** (0.0840)
Observations	116,323	8,532	107,791	81,549
Adjusted R-squared	0.364	0.261	0.352	0.364

Note: Significance levels: \* $<0.05$ , \*\* $<0.01$ . All specifications contain school fixed effects and are clustered at the school level.

**Average partial effects**

	<b>APE</b>	<b>St. Dev.</b>	<b>St. Error</b>	<b>observations</b>
interactions	0.1543322**	0.1543322	0.0049018	118941
interactions 03/05	0.1079994**	0.1079994	0.00031315	118941

the baseline variables (being black, age, gender, time dummies, *not shown*), education of the mother and of the child works indicator. The coefficients for tap water, electricity and income are not significant in the rural sample. This is an interesting result as the unconditional differences between the families with access to tap water and the families without access to tap water (table 3) lead us to believe that most of the variation is to be found in this sample. One possible explanation for this finding could be that families in rural areas have better ways how to cope with missing access to piped water whereas in urban areas safe substitutes are not easily available. Another explanation could be the uncertainty about the location of the access point and the source of the fresh water. The SAEB question does not allow to distinguish private in-house access to piped water from shared connections or connections on the plot of families. Additionally, Brazilian census data shows e.g. for the year 2000 that almost 60% of the rural households rely on water from wells or springs even if they have a piped water connections. Said differently, more than half of the households have access to piped water at home, but the source is not the *à priori* publicly provided network but some other source on their property. In urban areas, only 7% of the households in 2000 relied on water from own wells or springs. It would be interesting to take a closer look at these differences to see whether differences in water quality or access point drive the difference of results between the two samples. However our data does not allow to go deeper into these direction.

The last column of table 5 shows the results from the baseline specification run on a sample consisting only of the 2003 and 2005 wave. We exclude 2001 because of the concerns about data quality raised above.<sup>25</sup> The results are similar to those including the observations from the year 2001. However, the coefficient of tap water decreases by more than half of its magnitude and loses some significance. As most of the other coefficients remain the same in magnitude and significance, this results points to a decreasing significance of water access at home over time.

So far we have been including many different variables that capture at least part of unobserved abilities and economic success of the parents. As a last robustness check, we now add a large number of interaction terms to the specification in column 1 to allow for more functional flexibility. As we are only concerned about robustness of our results at this point, we won't report the coefficients of the interaction terms. We analyse effect heterogeneity of tap water access in the next section in greater detail. Here, we report the average partial effects of tap water for two different specifications at the bottom of table 5. The first specification includes interaction effects between ethnic background and drinking

---

<sup>25</sup>1999 drops earlier already because of the book variable which is not available for 1999.

water, sex and drinking water, age and education of the mother, age and drinking water, electricity and education of the mother, electricity and income, drinking water and income, child works and education of the mother, drinking water and child works, and all of the respective main effects. The average partial effect of drinking water in this specification is 0.1691 and is significantly different from zero at the one percentage level. When we focus on the years 2003 and 2005 (second bottom row in table 5), the average partial effect reduces to 0.1159, also significant at the one percentage level.<sup>26</sup> This confirms our findings from above and lends additional credibility to our results.

## 5 Heterogeneity of Effects

The literature has reported important complementarities between access to tap water and the education of the mother of a child (Gamper-Rabindran et al., 2010; Jalan and Ravallion, 2003). There are two beneficial effects of higher education of the mother. First, an educated mother probably knows better how to treat water to render it safe. Especially in Brasil, fresh water from the tap cannot be drunken as such but has to be cooked or filtered in allmost all regions of the country. In many cities tap water for drinking use is even replaced by bottled water. Second, hygiene is expected to be better with an educated mother as she probably trains her children better to follow basic hygienic rules. Her children don't attract water related diseases as often as children of non educated mothers. Another complementarity is reported for water access and income level. The literature disagrees with respect to the direction of the effects. On one hand, Jalan and Ravallion (2003) find with an Indian sample that only children that are not from extremely poor households benefit from access to tap water in terms of less diarrhea incidence and shorter duration of diarrhea. The authors suggest that resources and capabilities of uneducated and extremely poor families are too low, below a certain threshold, to make their children benefit from access to tap water. On the other hand, Gamper-Rabindran et al. (2010) find that Brazilian counties at lower development stages as compared to those at a higher stage of development experience larger reductions in infant mortality when access to tap water increases. Their results describe a substitution effect of income and water access at higher levels of development.

The above literature focuses on the effects of access to piped water on health outcomes of

---

<sup>26</sup>Note that we demeaned the income and the age variable in order to make the average effects comparable to the OLS results. In additional robustness checks, we included further interactions such as for example electricity times drinking water. The average partial effect of access to piped water at home on test scores always remains within the range of 0.11 and 0.18 and is always significantly different from zero.

children. As the relationship between access to tap water and schooling results acts most probably via the health status of the children and also shows long term effects of reduced health, we now investigate whether we find the reported effects also for our specification. Table 6 shows the results. We focus on the urban sample as we found no effects in the rural sample. We are interested in the effects of access to tap water on test scores at different income levels and also in the complementarities of human capital at home and tap water at home at different income levels of the families. That is why we consider three different intercepts. The first column is evaluated at mean income. These results are comparable to the OLS results above. The second and third column evaluate the same effects with the income reference group shifted one standard deviation below the mean and one standard error above the mean. In order to allow for different effects of education interacted with tap water access at different income levels, we interact water access with education and income.<sup>27</sup>

We evaluate all effects for families with mean income, below average and above average income. We aggregate higher education starting from the ninth grade on, as we don't expect effects to differ above a certain educational threshold. We don't find an effect of tap water on test scores of children of uneducated mothers. However, the main effect of tap water and the interaction effect of education level and tap water keeping income constant are jointly significant at the 1 percent level at mean income and at one standard deviation below mean income for all educational levels. That is there is a positive significant correlation between test scores and piped access to water at home for children that have a mother with at least primary education. At one standard deviation above the mean, only higher education and tap water are jointly significant and only at the ten percent level. This means that the effect of tap water is vanishing at higher income levels. The coefficients of the interaction effects increase with educational level at all three income levels. Testing for differences between the joint effects of tap water and education shows that only the effect of higher education interacted with tap water is significantly different from the effect of primary schooling interacted with tap water. For mean and lower income this holds at the one percent level, for higher income only at the five percent level. The joint effect of tap water and a maximum of eight years of schooling (secondary 1) is not significantly different from both other interaction effects. In addition to the increase of the effect with

---

<sup>27</sup>Note that the specifications we run are equal to evaluate the interaction effect of tap water, education and income at three different income levels. We shift the intercept by transforming the income variable. Thereby we constrain the slope to be equal at all three income levels, but the intercepts may differ. We are interested in the different intercepts: the additional effect of water at different education levels at different income sections.

education, the coefficients are largest at income one standard deviation below the mean and lowest for income one standard deviation above the mean.

To sum up, we find no effect of piped water access for children from households where the mother has no education at all. Children with an educated mother from mean income and below mean income families benefit from tap water access at home and the size of the effect increases with the educational level. If households dispose of higher income, tap water has no beneficial effect on children's test scores. These results allow interesting conclusions. If taken causally, this means that more time spent in school by mothers increases the return to access to tap water for children from poorer families but not for children from high income families. Stated differently, our results suggest that in richer families children don't benefit from tap water anymore because income substitutes for education. This is plausible as richer families have access to better health care and can afford additional classes for their children when they are ill of water related diseases. This makes the education of the mother less decisive. Also, richer families can afford expensive bottled water for drinking and cooking use. Thereby they reduce the probability of water related diseases as compared to low income families and that could be another reason why better education of the mother has smaller or no effects on the test scores of their children. Our results are with the evidence found by Gamper-Rabindran et al. (2010) who find that the effects of water are higher at lower levels of development. We also find that the effects increase with education (as Jalan and Ravallion (2003)) but don't find that this increase is conditional on some minimum income level.<sup>28</sup>

## 6 Conclusion

We analyze the effect of access to piped drinking water at home on educational achievements of fourth graders in Brazilian primary schools. We find that there is a positive and significant effect which explains around 11 percent of the standard deviation of average test scores. We estimate a school fixed effects estimator and thereby control for all time invariant variation across schools, as e.g. differences in quality of schools, which could explain variation in test scores. School effects also control for time invariant determinants of differences in test scores across municipalities. This is important as institutional quality of the municipalities probably impacts on schooling quality and the probability of access to piped water of the children in the same time. Further, we control for several proxies for the socio-economic background of the children's families as educational attainments of the

---

<sup>28</sup>(Jalan and Ravallion, 2003) report that they find effects starting for households above the 40% percentile of the income distribution.

Table 6: Effect Heterogeneity, Urban Sample

Dependent Variable: Test Score Math, Grade 4			
	mean inc	- 1 SD	+ 1 SD
Tap water	0.0304 (0.0911)	0.0918 (0.151)	-0.0310 (0.0654)
Income	-0.0816** (0.0364)	-0.0816** (0.0364)	-0.0816** (0.0364)
Income*Water	0.0319 (0.0377)	0.0319 (0.0377)	0.0319 (0.0377)
Mom_primary	0.103 (0.0962)	0.162 (0.163)	0.0431 (0.0643)
Mom_sec1	0.0970 (0.101)	0.168 (0.174)	0.0259 (0.0938)
Mom_higher	0.0734 (0.0993)	0.192 (0.160)	-0.0451 (0.0797)
Mom_don't know	0.0702 (0.0982)	0.254 (0.170)	-0.114* (0.0655)
Mom_primary*Income	0.0309 (0.0404)	0.0309 (0.0404)	0.0309 (0.0404)
Mom_sec1*Income	0.0369 (0.0502)	0.0369 (0.0502)	0.0369 (0.0502)
Mom_higher*Income	0.0616 (0.0409)	0.0616 (0.0409)	0.0616 (0.0409)
Mom_don't know*Income	0.0956** (0.0431)	0.0956** (0.0431)	0.0956** (0.0431)
Mom_primary*Water	0.0805 (0.0960)	0.139 (0.160)	0.0222 (0.0756)
Mom_sec1*Water	0.155 (0.102)	0.243 (0.177)	0.0682 (0.0946)
Mom_higher*Water	0.256** (0.101)	0.341** (0.158)	0.171* (0.0881)
Mom_don't know*Water	0.122 (0.0979)	0.112 (0.167)	0.133* (0.0774)
Mom_primary*Inc*Water	0.0302 (0.0419)	0.0302 (0.0419)	0.0302 (0.0419)
Mom_sec1*Inc*Water	0.0453 (0.0517)	0.0453 (0.0517)	0.0453 (0.0517)
Mom_higher*Inc*Water	0.0441 (0.0412)	0.0441 (0.0412)	0.0441 (0.0412)
Mom_don't know*Inc*Water	-0.00529 (0.0447)	-0.00529 (0.0447)	-0.00529 (0.0447)
Constant	0.406*** (0.115)	0.248 (0.167)	0.563*** (0.0907)
Observations	107,791	107,791	107,791
Adjusted R-squared	0.354	0.354	0.354

Note: Significance levels: \* $<0.1$ , \*\* $<0.05$ , \*\*\* $<0.01$ . All specifications contain school fixed effects and are clustered at the school level.



parents or the availability of electricity at home and a proxy for permanent income. Our estimate of the effect of piped water access on test scores remains robust to the inclusion of these variables and does not change in magnitude.

In accordance with evidence of heterogeneous effects of access to piped water on health outcomes of young children, our results show two different complementarities. First, the effect of access to piped water increases with the level of education of the mother. If the mother is uneducated there is no positive effect of water access. Better educated mothers may know themselves more about how to use water safely and may also better know to teach their children about hygiene. Further, it may be that children with access to tap water and therefore better health benefit more from the educational level of their parents. Second, we find that the joint effect of tap water and education of the mother is stronger at lower levels of income and the positive correlation between access to water and schooling achievement almost vanishes for children from families with permanent income one standard deviation above the mean. These results tell a plausible story as richer families may be able to provide better health care and additional classes to their children if they are ill, which may in part substitute for knowledge about hygiene and water related diseases, or piped water access.

Based on recent literature about the development of cognitive development and the decisive role that health in early childhood plays for the development of cognitive abilities, we suggest that our estimate depicts the positive effect of a better health environment in earlier childhood due to access to piped water. Piped water access, in combination with appropriate sanitation access, has been shown to sharply reduce the incidence of water related diseases (Gamper-Rabindran et al., 2010; Jalan and Ravallion, 2003) which mostly affect young children and have strong effects on their health. To the best of our knowledge we are the first to provide evidence for a link between piped water access and *qualitative* educational achievement as measured by test scores from standardized school exams. Our variable of interest, access to piped water, allows to capture the long term effect of water related diseases on human capital formation whereas the recent literature provides evidence of short term effects on *quantitative* measures of educational attainment as for example reduced absenteeism.

Various robustness checks with respect to the inclusion of controls, more flexibility in functional form and different sample restrictions and the fact that our variable of interest remains unaffected by these tests in almost all regressions, make us relatively sure that endogeneity is less of an issue in our study. However, in the end, we cannot be sure to control for all unobserved heterogeneity. As the policy implications of our results, if taken to

be causal, are relatively clear cut – focus infrastructure interventions on poor families and complement these efforts with educational support of the mothers and pro-poor subsidies to pay for uptake fees – we call for the collection of data sets allowing to focus on the effect of access to piped water on cognitive development to verify and further substantiate our results. This seems especially crucial as current policy practice mostly acts in the opposite way: Infrastructure expansion in developing countries normally follows income, i.e. richer areas are connected earlier to the network, and uptake fees for water connections are not subsidized whereas prices for already connected users are heavily subsidized.

## References

- Alderman, H. and Behrman, J. R. (2006). Reducing the Incidence of Low Birth Weight in Low-Income Countries Has Substantial Economic Benefits. *World Bank Research Observer*, 21(1):25–48.
- Alderman, H., Hoddinott, J., and Kinsey, B. (2006). Long Term Consequences of Early Childhood Malnutrition. *Oxford Economic Papers*, 58(3):450 – 474.
- Alderman, H., Hoogeveen, H., and Rossi, M. (2009). Preschool Nutrition and Subsequent Schooling Attainment: Longitudinal Evidence from Tanzania. *Economic Development and Cultural Change*, 57(2):239–260.
- Almond, D. and Currie, J. (2010). Human Capital Development Before Age Five. Working Paper 15827, National Bureau of Economic Research.
- Baird, S., Hicks, J. H., Kremer, M., and Miguel, E. (2012). Worms at Work: Long-run Impacts of Child Health Gains.
- Bleakley, H. (2007). Disease and Development: Evidence from Hookworm Eradication in the American South. *Quarterly Journal of Economics*, 122(1):73–117.
- Bobonis, G. J., Miguel, E., and Puri-Sharma, C. (2006). Anemia and School Participation. *Journal of Human Resources*, 41(4).
- Case, A. and Paxson, C. (2008). Stature and Status: Height, Ability, and Labor Market Outcomes. *Journal of Political Economy*, 116(3):499 – 532.
- Checkley, W., Buckley, G., Gilman, R. H., Assis, A. M., Guerrant, R. L., Morris, S. S., Mølbak, K., Valentiner-Branth, P., Lanata, C. F., Black, R. E., Malnutrition, T. C., and Network, I. (2008). Multi-Country Analysis of the Effects of Diarrhoea on Childhood Stunting. *International Journal of Epidemiology*, 37(4):816–830.
- Cunha, F. and Heckman, J. (2007). The Technology of Skill Formation. Working Paper 12840, National Bureau of Economic Research.
- Dillingham, R. and Guerrant, R. L. (2004). Childhood Stunting: Measuring and Stemming the Staggering Costs of Inadequate Water and Sanitation. *Lancet*, 363(9403):94 – 95.
- Fewtrell, L., Kaufmann, R. B., Kay, D., Enanoria, W., Haller, L., and Colford Jr, J. M. (2005). Water, Sanitation, and Hygiene Interventions to Reduce Diarrhoea in Less Developed Countries: A Systematic Review and Meta-Analysis. *Lancet Infectious Diseases*, 5(1):42 – 52.

- Fogel, R. W. (1994). Economic Growth, Population Theory, and Physiology: The Bearing of Long-Term Processes on the Making of Economic Policy. *American Economic Review*, 84(3):369–95.
- Gamper-Rabindran, S., Khan, S., and Timmins, C. (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.
- Glewwe, P. and Miguel, E. A. (2008). The Impact of Child Health and Nutrition on Education in Less Developed Countries. In Schultz, T. P. and Strauss, J. A., editors, *Handbook of Development Economics*, volume 4, chapter 56, pages 3561–3606. Elsevier, Amsterdam.
- Günther, I. and Fink, G. (2010). Water, Sanitation and Children’s Health: Evidence from 172 DHS Surveys. World Bank Policy Research Working Paper 5275, The World Bank.
- Guerrant, R. L., Kosek, M., Moore, S., Lorntz, B., Brantley, R., and Lima, A. A. M. (2002). Magnitude and Impact of Diarrheal Diseases. *Archives of Medical Research*, 33(4):351–355.
- IBGE (2008). Metodologia do Censo Demográfico 2000. Technical report, Instituto Brasileiro de Geografia e Estatística (IBGE).
- Jalan, J. and Ravallion, M. (2003). Does Piped Water Reduce Diarrhea for Children in Rural India? *Journal of Econometrics*, 112(1):153–173.
- Kremer, M., Leino, J., Miguel, E., and Zwane, A. P. (2011). Spring Cleaning: Rural Water Impacts, Valuation, and Property Rights Institutions. *Quarterly Journal of Economics*, 126(1):145 – 205.
- McKenzie, D. (2005). Measuring Inequality with Asset Indicators. *Journal of Population Economics*, 18:229–260.
- Miguel, E. and Kremer, M. (2004). Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities. *Econometrica*, 72(1):159–217.
- Moore, S., Lima, A., Conaway, M., Schorling, J., Soares, A., and Guerrant, R. (2001). Early Childhood Diarrhoea and Helminthiasis Associate with Long-term Linear Growth Faltering. *International Journal of Epidemiology*, 30(6):1457–1464.

- Prüss-Üstün, A., Robert Bos, F. G., and Bartram, J. (2008). Safer Water, Better Health: Costs, Benefits and Sustainability of Interventions to Protect and Promote Health. Technical report, WHO, Geneva.
- SAEB (2004). Sistema Nacional de Avaliação da Educação Básica – Microdados do Saeb 2001. Technical report, Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP), Brasília.
- SAEB (2006). Relatório Nacional Saeb 2003. Technical report, Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP), Brasília.
- SAEB (2008). Sistema Nacional de Avaliação da Educação Básica – Microdados do Saeb 2005. Technical report, Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP), Brasília.
- Thomas, D. and Strauss, J. (1997). Health and Wages: Evidence on Men and Women in Urban Brazil. *Journal of Econometrics*, 77(1):159–185.