The Effects of In-Utero Shocks on Cognitive Test Scores: Evidence from Droughts in India*

Bryce Millett
Harvard University
mmillett@fas.harvard.edu

Manisha Shah
University of California–Irvine and NBER
m.shah@uci.edu

January 2012

Abstract

We use cognitive test scores to investigate the impact of in-utero drought exposure on the human capital of children, exploiting fluctuations in monsoon rainfall over time and across districts in rural India. Our data is unique in that we have four distinct measures each for literacy and numeracy for close to 3 million children, who are either currently enrolled, have dropped out or never enrolled in school. We find that children exposed to drought in-utero score significantly worse on math and reading tests. Experiencing a drought in-utero is associated with being 2.6 percentage points less likely to recognize numbers from 1-10 and 1.2 percentage points less likely to do a simple subtraction problem (from a baseline percent of 53.9% and 61.6%, respectively). We find that the negative impacts on math are largest for in utero shocks, diminished for shocks before age two, and stop being negative at age two. While the magnitude of the effects on reading tend to be smaller, they are still negative and statistically significant. We then explore some potential mechanisms and provide evidence that these children are less likely to be on track in school and less likely to ever enrol.

JEL Codes: I2, I1, J1

*We would like to thank Ed Glaeser, Larry Katz, Michael Kremer, and Emily Oster for helpful comments.

1 Introduction

Recently there has been an upsurge of literature in economics suggesting that the in-utero period could be one of the most critical periods in a person's life, shaping future abilities and health outcomes as well as later life economic outcomes such as earnings (Almond and Currie, 2011; Cunha and Heckman, 2008). Most of these studies are related to the fetal origins hypothesis (Barker, 1994) which states that economic and environmental conditions during pregnancy may have long-term impacts on health and socioeconomic status (see e.g., Almond (2006); Black et al. (2007); Dêschenes et al. (2009); Royer (2009)). This is because the intrauterine environment and nutrition in particular can impact the fetus' metabolism, which can lead to future diseases like obesity, diabetes and/or cardiovascular problems (see Almond and Currie (2011) for a more in-depth discussion).

In this paper we look directly at measures of educational achievement, and investigate the impact of in-utero drought exposure on cognitive ability, exploiting fluctuations in monsoon rainfall over time and across districts. To date, much of this fetal origins literature has focused on health outcomes (Currie et al. (2010); Almond and Mazumder (forthcoming); Hoddinott and Kinsey (2001); Lawlor et al. (2006); Kudamatsu et al. (2010)) and less on educational outcomes. The research which has focussed on education generally uses educational attainment and enrollment as outcomes of interest (e.g. Fung (2010); Maccini and Yang (2009); Alderman et al. (2009)). In this paper we are able to focus directly on cognitive ability as we have various test outcomes for each child in the data.

While there are some recent working papers which look at similar cognitive test outcomes in children, notably Aguilar and Vicarelli (2011) and Akresh et al. (2010); one advantage of this paper is that the data are much richer. We observe nearly 3 million children in 20 cohorts in almost every state in India. We have four distinct measures of math achievement and four distinct measures for reading for each child regardless of whether he is currently enrolled or not. Since the survey was conducted every year over five years, we can control for age, year of survey, and district. In addition, our data allow us to look at more standard educational measures such as school enrollment, drop-out behavior, and being on track in school (age for grade).

We find that children exposed to drought in-utero score significantly worse on math and reading tests. For example, experiencing a drought in-utero is associated with being 2.6 percentage points less likely to recognize numbers from 1-10 and 1.2 percentage points less likely to do a simple subtraction problem (from a baseline percent of 53.9% and 61.6%, respectively). While the magnitude of the effects on reading tend to be smaller than the effects on math, they are still negative and significant. We explore the timing of these shocks and find that drought exposure in-utero matters the most. The negative impacts persist at ages one and two, though not as large in magnitude and move to zero by age three. We also show that these results are robust to concerns related to selection on fertility (i.e women are not responding to droughts by delaying fertility) and migration.

We then explore some of the potential mechanisms that might explain these results. While we estimate well-identified impacts of in-utero drought exposure on achievement and schooling outcomes, the discussion on mechanisms is more suggestive given we cannot randomly assign schooling to these children. There are a few relevant channels to explore in terms of why in-utero drought exposure could affect cognitive ability outcomes.

First, children exposed to in-utero drought might simply get less schooling. The less schooling could be a result of biological or social factors. For example, the fetal origins hypothesis links in-utero nutrient deficiencies to worse health outcomes and decreased cognitive functioning. Because these drought exposed children will be less healthy on average, this could impact school attendance. In fact, Maccini and Yang (2009) show that Indonesian women who experienced higher than normal rainfall in their birth year attend more years of schooling precisely because they tend to be healthier.[1] Less schooling could also stem from

[1]Unfortunately we do not have health outcomes in the ASER data so we cannot test whether these children are less healthy and are attending school less as in Maccini and Yang (2009). However, using another dataset of young children in India (the NFHS-2) which measures children's height and weight, we show that children who were exposed to drought in-utero are significantly smaller both in terms of height and weight. This alludes to worse health outcomes for Indian children exposed to drought in-utero suggesting they might be less likely to attend school. See Section 5.2.

social factors. For example, parents may simply respond to lower ability children by pulling them out of school or not enrolling them at all.

A second channel could be purely biological in that drought exposed children are simply lower ability, so conditional on the same level of schooling, they will still score lower on achievement tests because they learn less efficiently.

We explore some of these mechanisms by investigating the effect of drought on school enrollment, dropping out behavior, and being on track in school (given age). We find that children exposed to in-utero drought are less likely to be enrolled in school suggesting these children get less schooling overall. Interestingly, there are no significant differences by gender. In addition, we find children exposed to drought in-utero are less likely to be on track in school given their age, again alluding to the less schooling hypothesis. These results also suggest they might be held back more because they learn less efficiently in school. Unfortunately our data do not allow us to disentangle which part of the in-utero effect is purely "biological" and which part is parental or teacher response.

2 Data and Sample

2.1 Aser Data

Every year since 2005, the NGO Pratham has facilitated an innovative exercise for India: that of implementing the Annual Status of Education Report (ASER). We use ASER data in this paper. The core of the survey is that simple tools are used to assess children's ability to read simple text and do basic arithmetic. Within the span of a hundred days, the survey reaches *every* rural district in the country: over 570 districts, 15,000 villages, 300,000 households and 700,000 children. ASER is the largest annual data collection effort with children in India. It is also the only annual source of information regarding learning levels of children in elementary school. We have data on children for every year since 2005 inclusive of 2009, giving us a sample size of over 3 million rural children. The sampling strategy used generates a representative picture of each district, and all rural districts are surveyed. Though the data is collected annually, each year is cross-sectional.[2]

What differentiates ASER data from other educational datasets is that every child within a household is tested. This means we have cognitive ability test score outcomes for children who are currently enrolled in school, who have never enrolled in school, and who have dropped out of school. Generally test scores only tend to be available for children who are currently enrolled in school, not for children who have dropped out or never enrolled. In addition, ASER data is quite unique in terms of the large sample size. Normally when surveys involve testing, sample sizes are limited due to high costs of testing children. However, aggregating across the years of ASER data gives us a sample size of over 3 million rural children.

In Table 1 we describe the characteristics of the children in our sample as well as their test scores. The average age is about nine and a half, and the average grade (standard) is 4.6. The sex ratio is somewhat skewed, with 54.4% boys. This skewed sexratio is expected given this is Indian data where sex selected abortion and female infanticide are common and girls have become relatively more scarce over time, especially in rural India (Sen, 1992; Jha et al., 2006).

The ASER surveyors ask each child four questions each in math, reading (in their native language), and English reading. The four math questions are whether the child can recognize numbers 1-9, recognize numbers 10-99, subtract, and divide. The scores are coded as 1 if the child correctly answers the question, and 0 otherwise. In addition, there are four literacy questions: whether the child can recognize letters, recognize words, read a paragraph, and read a story. We use the scores for numeracy and native-language literacy in our analysis, and those scores are reported in Table 1. In addition, we calculate a "math score" variable, which is the sum of the scores of the four numeracy questions. For example, if a child correctly recognizes numbers between 1-9 and 10-99, and correctly answers the subtraction question, but cannot correctly answer the division question, then that child's math score would be coded as 3. The "reading score" variable is calculated in exactly the same way.

[2]For more information on ASER, see `http://www.asercentre.org/ngo-education-india.php?p=ASER+survey`

Approximately 65% of the children tested can recognize numbers between 1 and 9, and about 38% can correctly do a division problem. The reading scores are slightly higher: nearly 90% of children tested can recognize letters and 45% can read a story. The ASER data also contains scores on English reading, but we assume that most of the variation in these scores will be caused by whether the child learns English at school, rather than cognitive ability, so we exclude it from our analysis.

2.2 Rainfall Data

Drought has affected millions of people over the past two centuries in India. The rain that comes during the monsoon season (June–October) is essential to raising crops, and if rainfall is below normal levels, many regions experience deleterious living conditions since most of the rural poor depend on rain-fed agriculture. In fact, almost 70 percent of the total net area sown in India is rainfed; and 66.2 percent of rural males and 81.6 percent of rural females report agriculture (as cultivators or laborers) as their principal economic activity (Mahajan and Gupta, 2011). While there is plenty of evidence showing droughts adversely affect agricultural output in India, Pathania (2007) does a nice job of empirically showing the close correlation between rainfall fluctuations and agricultural output at the all India level. For example, the annual yields of wheat and rice, the two major cereal crops consumed, are clearly affected adversely by droughts; a second consecutive year of drought leads to a 9.4% decrease in wheat and 20% in the case of rice (Pathania, 2007). Given the majority of rural residents are engaged in agricultural activities in India, and that the majority of this land is rainfed, it is clear that drought adversely affects pre- and post-natal women, children, and other household members.[3]

To determine drought years and districts, we use monthly rainfall data which is collected by the University of Delaware.[4] The data covers all of India in the period between 1900-

[3]While there are many potential channels through which this can occur, the point of this paper is not to identify the channels. However, some potential mechanisms could be income effects, maternal stress, labor and time use changes, etc.

[4]The data is available at:`http://climate.geog.udel.edu/~climate/html_pages/download.html#`

2008. The data is gridded by longitude and latitude lines, so to match these to districts, we simply use the closest point on the grid to the center of the district, and assign that level of rainfall to the district for each year. To define a drought, we use the cutoff of the Indian Meteorological Department, which is monsoon rainfall that is 75% of the 10-year average of rainfall for that district. Figure 1 shows the number of districts in each year with a drought, from a low of 5.6% of districts in 1998 to a high of 47.4% of districts in 2002. Figure 1 also indicates that there is both a lot of variation over time and across districts in terms of drought exposure.

In Figure 3 and Figure 4 we graph the relationship between each child's math score and in-utero drought exposure and each child's reading score and in-utero drought exposure, respectively. The figures allude to the relationship we will find in the empirical results: children who were exposed to drought in-utero are more likely to score lower on the math and reading tests. Whether this relationship holds once we control for various observable characteristics is what we test below in Section 3.1

3 Empirical Strategy and Timing of Birth

Unfortunately we do not have the exact date of birth for the children in our sample, only the current age, so we need to calculate which drought years will affect the in-utero environment. Since we observe child age and year of survey, we assume that each child has already had his/her birthday that year, that is, that his/her year of birth is the year of the survey minus age (e.g. a child who reports being 10 in the 2008 survey is coded as being born in 2008-10=1998). Since the Indian Monsoon (and, thus, droughts) typically takes place in the summer, and harvests happen the following fall, it is likely that children born in a given year will be more affected by a drought in the year before their birth than one in the year they are born. In Figure 2, we show this more explicitly. We assume a child born in 2008, on average, is born June 30. This child will be in-utero roughly from October 2007 to June

P2009

2008. Thus, the harvest in 2007 (which is in turn affected by the 2007 rains) is likely to be the biggest determinant of his nutritional environment in-utero (as opposed to harvest 2006 or 2008). Of course, there will be children in the sample whose age is recorded incorrectly, and those who are born at the very end or beginning of the year for whom other drought years might be more relevant. However, unless these errors are systematic, they will simply add noise to our estimation and attenuate the results.

Our main analysis is very straightforward. We take advantage of the quasi-random nature of droughts within district (and across districts within a year) in order to measure the effect of drought on test scores. We estimate the following regression:

$$S_{ijt} = \alpha + \beta\delta_{j,t-1} + \gamma_j + \phi_t + \epsilon_{ijt} \quad (1)$$

where S_{ijt} is the test score of student i in district j born in year t, δ_{jt} is an indicator for whether there was a drought in district j in year $t-1$, γ_j is a vector of district fixed effects, ϕ_t is a vector of year-of-birth fixed effects. This way, any unobservables that vary with district or year will be absorbed by the fixed effects, and thus the effect we pick up will be that of being born in-utero during a drought. β is our coefficient of interest and it is the causal impact of in-utero drought on the various test scores (or cognitive ability). Standard errors are clustered at the district level. We discuss some potential selection issues in Section 4 below.

One possible issue with using droughts as quasi-random shocks, is that it is possible they are correlated over time. There are certainly districts in which droughts are more common in all years, but this should not affect our empirical results, since the district fixed effects imply we are using within-district variation in timing of droughts to identify causal effects. However, if it is the case that droughts this year are correlated with droughts next year, then this undermines the effectiveness of our cohort identification as it will be difficult to separate the effects of droughts in the year before birth with those in early life (or before pregnancy). This will also bias our results downward, since many of the "control" children will have been in these early life stages during droughts, and the effects on them could be significant. We test for serial correlation directly in Appendix Table A1.

Recent papers using Indian rainfall data have not found evidence of serial correlation in rainfall data (see for example, Kaur (2011); Pathania (2007)). Nevertheless, Table A1 estimates the autocorrelation in rainfall in our data. An observation is a district-year. The dependent variable in both regressions is the deviation from mean rainfall in the current year (in inches), where deviation is simply defined as current year rainfall minus the mean rainfall in sample period. The independent variable is deviation from mean rainfall last year, constructed in the same way. The specification in column 2 contains year fixed effects, while column 1 does not. These specifications follow closely the methodology used in Kaur (2011). In column 1 we find no significant evidence of serial correlation. In column 2 once we include year fixed effects, the coefficient becomes negative and statistically significant, however, the magnitude of the effect is close to 0. Therefore, it has very little economic significance.

3.1 Main Results: Cognitive Test Scores

Table 2 presents our main estimates of the effect of in-utero droughts on cognitive test scores. Panel A shows the effect of in-utero drought exposure on the four math questions as well as the math score variable, which is the simple sum of the four questions. The effect on the overall math score is negative and significant, and represents about 8.9% of one question. The negative effects are quite a bit higher for the number recognition questions relative to the subtraction and division questions. Experiencing a drought in-utero is associated with being 2.6 percentage points less likely to recognize numbers from 1-10 and 1.2 percentage points less likely to be able to do a simple subtraction problem (from a baseline percent of 53.9% and 61.6%, respectively). The magnitude of the effect for recognizing numbers 1-9 is smaller, and while the coefficient on division is negative, it is not statistically significant.

Panel B of Table 2 shows the effect of in-utero drought exposure on the reading questions.

The two more difficult questions—reading a paragraph and reading a story—are statistically significant at the .01 level. Children in-utero during a drought are 0.7 percentage points less likely to be able to read a paragraph and 0.7 percentage points less likely to be able to read a story (from a baseline percent of 60.8% and 44.6%, respectively). The effect on the overall reading score is negative (-0.9 points), but not statistically significant. While the magnitude of the effects on reading tend to be smaller than the effects on math, they are still negative.

Panels C and D of Table 2 show the results using household, rather than district, fixed effects. Overall, the results are remarkably similar. This is a little surprising, given we are now identifying off within-household differences in exposure to in utero drought, and one might expect the differences to be larger between rather than within households. However, the results lend credence to our assumption that droughts are quasi-random, and that there are no systematic ex-ante differences between those who are exposed to drought in utero and those who are not, once we control for cohort and district effects.

3.1.1 Main Results: Gender and Timing Issues

The main results are strikingly similar when broken down by gender. Panels A and B of Table 3 show reading and math scores for boys, and Panels C and D show the same effects for girls. The effects on reading are slightly higher for girls, though not significantly so. The impacts on math are higher for boys, and this is a statistically significant difference. These results are in contrast to Maccini and Yang (2009), who find that early-life rainfall in Indonesia affected only females (not males) later in life. Their effect is strongest for rainfall in infancy, and they hypothesize that parents with reduced income and resources tend to take from their daughters rather than their sons. When we closely investigate timing of the shock (see Table 4) we find that the strongest effect of drought on the math score is during the in-utero period. Therefore, in our case, we find a much stronger effect for droughts in the year before birth.

In Table 4, we regress the test score on experiencing drought one year before the in-utero period up to age two. The results indicate that the magnitude of the effects of drought decrease after the in-utero period and stop mattering at age two.[5] Interestingly, the magnitudes of the in-utero effects are similar to the main results even after we add these additional years of drought in the regressions. Our results are very similar to Akresh et al. (2010) who also find that negative education impacts are largest for in utero shocks, diminished for shocks before age two, and have no impact for shocks after age two in rural Burkina Faso. Unlike Aguilar and Vicarelli (2011) who find that negative impacts on child development persist four to five years after rainfall shocks in rural Mexico, we do not find any evidence that the negative impacts persist after age two.

3.1.2 Alternative Rainfall Measures

So far we have investigated the impact of drought on test scores since drought is the most serious weather shock that rural households in rainfed agricultural areas face. However, it is possible that other measures of rainfall might also be relevant weather shocks. For example, Maccini and Yang (2009) find that *higher* early-life rainfall leads to improved health, schooling, and socioeconomic status for women. In Table 5 we measure the effect of two different measures of rainfall on test scores. The dependent variable in both specifications is math score, defined previously as the sum of questions answered correctly out of the four numeracy questions given in the ASER survey. The independent variable in column 1 is rainfall (in inches). The independent variables in column 2 are quintiles of rainfall (the third quintile, the middle, is omitted). Both specifications include district, age, and year of survey fixed effects and standard errors are reported in parentheses.

Column 1 indicates that a one standard deviation increase in rainfall (788.5 inches) increases the math score by .078 points. This translates to a 4 percent increase off of a mean of 2.19. The quintile analysis is even more interesting in that the results are monotonic.

[5] In alternate specifications, we investigate the impact of later years of drought on test scores, but none of the effects are ever negative. In fact, current-year droughts seem to increase test scores in our data, which could be explained by lower wages making school enrollment and attendance more attractive in drought years (results available upon request).

The bottom two quintiles of rainfall are negatively associated with math test scores (relative to the middle quintile), and the top two quintiles are positively associated with test scores (relative to the middle). These results are significant at the .01 level. This is exactly what we would expect in a region where economic livelihood depends on the "right" amount of rainfall.

4 Potential Selection Bias Issues

4.1 Migration

One weakness with our data is that we do not have information on location of birth. Therefore we assume that the current district is the same as the birth district when we assign each child the district level drought measure which corresponds to his/her year of birth. Because these children are relatively young, ages 3-15 with a mean age of 9, we do not think this is a strong assumption. Nevertheless, we explore this issue below. If it is the case that higher ability families are more likely to migrate out of rural areas (say to urban areas), this will affect our empirical results. However in this case, our results will be lower bounds.

Various pieces of evidence point to statistics that out-migration rates are low for rural Indian families. For example, Topalova (2005) using data from the National Sample Surveys finds that only 3.6 percent of the rural population in 1999-2000 reported changing districts in the previous 10 years. Munshi and Rosenzweig (2009) using the Rural Economic Development Survey also conclude that rural emigration rates are low and actually declined between 1982 and 1999 in India. Pathania (2007) using Indian Census data also finds that only a small fraction of rural women reside in districts different from their district of birth. Using data from the 2001 census on internal migration, he shows that 82.65 percent of rural women aged 15-59 are residing in the district of birth and 96 percent of this demographic group is residing in the state of birth. He writes that marriage and the subsequent move to the house of the husband's family is the major reason for female migration, and most marriages are local. Given women have children after marriage, it is unlikely that many of our rural sample of women is moving after they have had children. While temporary migration of rural men in search of employment is more common, this will not affect our results.

To think that migration might bias our coefficient upward, one would have to imagine not only that lower ability kids move into drought-prone districts (which could be true), and that this effect is also large enough to affect the coefficient (unlikely, given the limited movement), but that the lower ability kids who are moving in are particularly likely to be certain ages, namely those which correspond to being born in a drought year. This seems extremely unlikely, especially since consistent district effects are differenced out in the analysis.

4.2 Selective Mortality and/or Fertility

One potential concern with trying to understand the effect of drought on cognitive development is that we only observe children who survive and make it into the sample; drought might impact child survival so we only observe children who have survived droughts changing the composition of children. This selection would most likely bias our results downward; since these are the children who survived, they are positively selected and probably do better on health and educational outcomes relative to the children who died off. Therefore, we are less concerned about bias from selective mortality.

However, another potential concern with our results could be if women are delaying and/or changing fertility patterns in response to droughts. For example, mothers may choose to wait out a drought year before having a child. Rural fathers could migrate during drought years in search of work and their absence would result in delayed fertility. If droughts are in fact impacting fertility decisions, the empirical results will most likely be biased upward, since the children being born in drought years would be negatively selected.

Since our dataset includes only children aged 3-16, both of these selection effects would show up as fewer children observed who were in-utero during a drought year (assuming that most of the selective mortality happens before age 3). In Table A2 in the Appendix we test whether this is the case. In column 1 of Table A2 we regress each child's year of birth on drought in-utero ($t - 1$) and drought the year before ($t - 2$). All specifications include district fixed effects and are clustered at the district level. We find that neither of the years of drought is a statistically significant predictor of birth year. The coefficient on ($t - 2$) is probably most relevant for selective fertility, as this is when fertility decisions are being made. In column 2 we go back even further to 4 years before birth, and again find that drought has no significant impact on the timing of birth regardless of the year prior to birth. Therefore, it appears that bias resulting from fewer children (and thus, a differently selected sample) during drought years is not a problem for our analysis.[6]

Another piece of evidence which points against selective fertility and selective migration are the household fixed effects results in Panels C and D of Table 2. If either selective migration or selective fertility is driving the results, then within-household variation in drought exposure should not affect cognitive test scores. This story relies on between household variation–i.e. that "good" households are acting differently with respect to droughts compared to "bad" households. That is, if "good households" are leaving the area after droughts, or delaying their fertility when there are droughts, then our sample of exposed children would be more heavily weighted toward "bad households", which could bias our results upward. However, the results with and without household fixed effects are extremely similar (if anything reading scores coefficients are a bit higher), which leads us to conclude that selection of this type is not contributing significantly to our estimates.

5 Mechanisms

5.1 Schooling Outcomes

It is reasonable to hypothesize that exposure to in-utero drought could affect schooling outcomes. There are two main channels through which this could occur. First, children

[6]Of course, this is not to say that there is not selective fertility or mortality going on—indeed, we see evidence of the skewed sex ratios that are indicative of sex-selective abortion and preferential treatment of male infants–it is just to say that this selection is not correlated with the timing of droughts, so that it will not affect our empirical results.

exposed to drought in-utero could be less healthy, which directly affects their ability to attend school. Second, droughts could directly affect cognitive ability, and parents could react by altering the amount of schooling given to these lower-ability children. If ability and schooling are substitutes—this would be reasonable in a setting in which, say, there are high returns to basic literacy and numeracy and little else—then parents might be more likely to enrol their lower-ability children, and less likely to pull them out of school at a young age. By contrast, if ability and schooling are complements—if high ability children get more out of each year, or have a higher probability of leaving the village for high-paying work—then parents would, on average, provide less schooling for their lower-ability children.[7]

Of course, we will not be able to directly test for the difference between these two possibilities. However, if the effect of in-utero drought on schooling is positive, we could speculate that the second effect dominates, and that schooling and ability are substitutes, but if the effect is negative we cannot distinguish between the "sickly child" mechanism and the ability-schooling complementarity mechanism, since we do not have random assignment of schooling. We can, however, test for the reduced-form effect of droughts on schooling, and in the following section we will provide some suggestive evidence on the mechanisms outlined above. We do find that the effect is negative suggesting that ability and schooling are complements.

Our empirical specification in this section is very similar to that outlined in the previous section, but instead of test scores, our left-hand-side variable is schooling outcomes. Specifically, we estimate the following regression:

$$Y_{ijt} = \alpha + \beta\delta_{j,t-1} + \gamma_j + \phi_t + \quad_{ijt} \tag{2}$$

which is the same as equation (1) above, but now the outcome variable is Y_{ijt} which is a schooling outcome binary variable: whether the child is reported as having dropped out,

[7]For empirical evidence of parents reinforcing initial cognitive ability differences across children by investing in higher ability children, see Ayalew (2005); Akresh et al. (2010); Frijters et al. (2010). In fact, the majority of recent empirical evidence suggests that parents reinforce initial cognitive differences and invest more in high ability children.

never enrolled, or is "on track". The last variable, "on track" is created to capture whether the student is progressing normally through the grades, and is coded as one if the child's grade is no less than his age minus six—that is, if a child is eight, he must be at least in the second grade, nine, at least in third grade, etc.—and zero otherwise. This is a way of capturing how children are doing in school.

Table 6 shows the effect of drought in-utero on schooling achievement. The first row shows the effects of in-utero drought on dropping out, which is positive but not statistically significant. The second row shows the effect of in-utero drought on never being enrolled in school, which is positive and significant—children born the year after a drought are 2 percentage points more likely to never enrol in school. This is quite a large effect given the mean of never enrolled is 2.8%. This is some evidence toward the hypothesis that schooling and ability are complements. If a child has some cognitive deficiency because of being malnourished in-utero, his parents are less likely to send him to school, presumably because schooling is less valuable for him/the family.

The third row of Table 6 shows the effect of in-utero droughts on being on track in school. The effect is negative and significant—children who are born in the year following a drought are 2.2 percentage points less likely to be on track, from a baseline of 81.3%. Panels B and C of Table 6 show the results separately for boys and girls. As with the test score outcomes, these results look fairly similar for boys and girls. One difference is that boys look to be somewhat less likely to be "on track" when exposed to drought in utero than their female counterparts. Whether this is due to differences in treatment effects, parental investment decisions, or simply differences in likelihood of being on the margin of passing we cannot say.

5.2 Child Health and In-Utero Drought Exposure

Though we cannot investigate the relationship between in-utero drought and health outcomes for the ASER sample children (since they do not ask about health outcomes), we turn to another dataset, the National Family Health Survey-2 (NFHS-2)[8] to investigate this relationship. This will allow us to investigate whether children born exposed to in-utero drought are less healthy and then are more likely to be absent from school. Though there is already evidence of this from other countries (see for example, Maccini and Yang (2009)), we would like to say something about Indian children using our rainfall data. We use the 1998-99 NFHS-2 India survey because this is the latest year that district identifiers are publicly available. We merge the rainfall data used above (which is at the district level) to the NFHS-2.

The NFHS-2 survey covers a nationally representative sample of more than 90,000 eligible women age 15-49 from 26 states that comprise more than 99 percent of India's population. The survey provides information on fertility, mortality, family planning, and important aspects of nutrition, health, and health care. The NFHS-2 measured children's (under 5) height and weight. Height and weight are a widely used proxy for overall health status and correlate positively with economic outcomes. For example, Case and Paxson (2008) show that height is positively correlated with earnings in the developed world. Similar patterns between height and wages for individuals in Brazil (Strauss and Thomas, 1998) and other developing countries have been shown (Behrman and Deolalikar, 1989). Similarly, underweight is correlated with future health problems and worse schooling outcomes.

In Table 7 we regress the dependent variables height (Panel A) and weight (Panel B) on in-utero drought exposure. We include birth year fixed effects, district and state fixed effects; all specifications are clustered at the district level. By including birth year fixed effects, in some sense we are investigating the relationship between height-for-age and weight-for-age and drought exposure of these young children. In addition, the NFHS-2 asks mothers for exact birth year, so we do not have to make any assumptions about age at the time of the survey when we merge the rainfall data.

The results are quite striking: in-utero drought exposure significantly decreases current

[8]This is also known as the Demographic and Health Survey (DHS) for India.

height and weight. For example, a child exposed to in-utero drought is 7.3 centimeters shorter on average than a similar child born during a non-drought year. While a small effect, it is statistically significant at .05 level and is about 6% of a standard deviation. In addition, children exposed to in-utero drought weigh 1.03 kilograms less on average, which is about 3% of a standard deviation.

In additional specifications (results available upon request from authors), we also include if there was a drought during the child's birth year. While the coefficients on in-utero drought exposure do not change in magnitude or significance levels, the coefficient on drought during birth year is also negative for both height and weight. However it is much smaller in magnitude and not statistically significant. Therefore, in-utero drought exposure matters more both economically and statistically than drought exposure during year of birth.

These results allude to the fact that the ASER children who were exposed to in-utero drought are not only cognitively disadvantaged but most likely have worse health outcomes as well. Earlier we discussed that the two most likely channels for the lower cognitive abilities due to drought are biological and health related. First, since these children are more likely to be exposed to hunger, malnutrition, and possibly famine-like conditions as a result of drought this may result in decreased cognitive functioning. The second channel through which in-utero drought exposure could affect cognitive ability is via health. Since it is well-established that children exposed to drought are less healthy, this could impact school attendance. Children who attend less school will most likely experience lower cognitive development. For the ASER children, it appears that both mechanisms are most likely at play.

6 Conclusion

This paper relates to the literature on the fetal origins hypothesis and the importance of "critical periods" in human capital acquisition in the economics literature. We show that children who are exposed to droughts while in utero score significantly worse on literacy and numeracy tests than their peers. The magnitude of the effects are strongest during the in utero period and stop being negative after the age of two. Further, we show evidence that these children are less likely to be on track in school and less likely to ever enrol. Interestingly, there are no large gender differences in the results. We argue that the results are causal and not due to differences in the sample of children exposed to these shocks. The size and scope of our data allows us to be confident about our conclusions and the precision of the estimates.

These results provide support for drought relief programs in developing countries, as well as those which encourage maternal nutrition during pregnancy. The results on schooling suggest that even if the goal of universal primary schooling is achieved, children in utero during droughts could remain behind their peers in skill acquisition.

References

Aguilar, Arturo and Marta Vicarelli, "El Nino and Mexican Children: Medium-Term Effects of Early-Life Weather Shocks on Cognitive and Health Outcomes," *mimeo*, 2011.

Akresh, Richard, Emilie Bagby, Damien de Walque, and Harounan Kazianga, "Child Ability and Household Human Capital Investment Decisions in Burkina Faso," 2010. IZA Discussion Paper 5326.

Alderman, Harold, Hans Hoogeveen, and Mariacristina Rossi, "Preschool Nutrition and Subsequent Schooling Attainment: Longitudinal Evidence from Tanzania," *Economic Development and Cultural Change*, January 2009, *57* (2), 239.

Almond, Douglas, "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*, 2006, *114* (4), 672712.

_ **and Bhashkar Mazumder**, "Health Capital and the Prenatal Environment: The Effect of Ramadan Observance during Pregnancy," *American Economic Journal: Applied Economics*, forthcoming.

_ **and Janet Currie**, "Killing Me Softly: The Fetal Origins Hypothesis," *Journal of Economic Perspectives*, Summer 2011, *25* (3), 153–172.

Ayalew, Tekabe, "Parental Preference, Heterogeneity, and Human Capital Inequality," *Economic Development and Cultural Change*, 2005, *53* (2), 381–407.

Barker, David J. P., *Mothers, Babies, and Disease in Later Life*, London: BMJ Publishing Group, 1994.

Behrman, Jere and Anil Deolalikar, *Causes and implications of seasonal variability in household food security*, The Johns Hopkins University Press, 1989.

Black, Sandra E., Paul J. Devereux, and Kjell G. Salvanes, "From the Cradle to the Labor Market? The Effect of Birth Weight on Adult Outcomes," *Quarterly Journal of Economics*, 2007, *122* (1), 40939.

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

Cunha, Flavio and James J. Heckman, "Formulating, Identifying and Estimating the Technology of Cognitive and Noncognitive Skill Formation," *Journal of Human Resources*, 2008, *43* (4), 738–82.

Currie, Janet, Mark Stabile, Phongsack Manivong, and Leslie L. Roos, "Child Health and Young Adult Outcomes," *Journal of Human Resources*, 2010, *45* (3), 518–547.

Dêschenes, Olivier, Michael Greenstone, and Jonathan Guryan, "Climate Change and Birth Weight," *American Economic Review Papers & Proceedings*, May 2009, *99* (2), 211–217.

Frijters, Paul, David W. Johnston, Manisha Shah, and Michael A. Shields, "Intrahousehold Resource Allocation: Do US Parents Reduce or Reinforce Child Ability Gaps?," August 2010. IZA Discussion Paper 5153.

Fung, Winnie Wan-Yi, "Malnutrition, infectious disease and economic development," 2010. Harvard University.

Hoddinott, John and Bill Kinsey, "Child growth in the time of drought," *Oxford Bulletin of Economics and Statistics*, September 2001, *63* (4), 409.

Jha, Prabhat, Rajesh Kumar, Priya Vasa, Neeraj Dhingra, Deva Thiruchelvam, and Rahim Moineddin, "Low male-to-female sex ratio of children born in India: National Survey of 1.1 Million Households," *Lancet*, 2006, *367*, 211–218.

Kaur, Supreet, "Nominal Wage Rigidity in Village Labor Markets," 2011. Harvard University Working Paper.

Kudamatsu, Masayuki, Torsten Persson, and David Strmberg, "Weather and Infant Mortality in Africa," September 2010. IIES, Stockholm University Working Paper.

Lawlor, Debbie, George Davey Smith, Richard Mitchell, and Shah Ebrahim, "Adult Blood Pressure and Climate Conditions in Infancy: A Test of the Hypothesis that Dehydration in Infancy Is Associated with Higher Adult Blood Pressure," *American Journal of Epidemiology*, April 2006, *163* (7), 608.

Maccini, Sharon and Dean Yang, "Under the Weather: Health, Schooling, and Economic Consequences of Early-Life Rainfall," *American Economic Review*, June 2009, *99* (3), 1006–26.

Mahajan, Vijay and Rajeev Kumar Gupta, "Non Farm Opportunities for Smallholder Agriculture," January 2011. Paper presented at the IFAD Conference on New Directions for Smallholder Agriculture.

Munshi, Kaivan and Mark Rosenzweig, "Why is Mobility in India so Low? Social Insurance, Inequality and Growth," 2009. NBER Working Paper No. 14850.

Pathania, Vikram, "The Long Run Impact of Drought at Birth on Height of Women in Rural India," November 2007. Working Paper.

Royer, Heather, "Separated at Girth: US Twin Estimates of the Effects of Birth Weight," *American Economic Journal: Applied Economics*, 2009, *1* (1), 49–85.

Sen, Amartya, "Missing Women," *British Medical Journal*, 1992, *304*, 587–588.

Strauss, John and Duncan Thomas, "Health, Nutrition and Economic Development," *Journal of Economic Literature*, June 1998, *36* (2), 766–817.

Topalova, Petia, "Trade Liberalization, Poverty, and Inequality: Evidence from Indian Districts," 2005. NBER Working Paper No. 11614.

Table 1: ASER Summary Statistics

	Mean	Std. Dev.	Observations
Child Characteristics			
Male	.545	.498	3,208,990
Age	9.53	3.62	3,229,037
Class	4.60	2.85	2,619,202
Math Scores			
	Mean	Std. Dev.	Observations
Can Recognize Numbers 1-9	.651	.477	2,499,352
Can Recognize Numbers 10-99	.539	.498	2,499,352
Can Subtract	.616	.486	2,499,352
Can Divide	.384	.487	2,499,352
Math Score	2.19	1.35	2,499,352
Reading Scores			
	Mean	Std. Dev.	Observations
Can Read Letters	.897	.304	2,729,313
Can Read Words	.754	.431	2,729,313
Can Read Paragraph	.608	.488	2,729,313
Can Read Story	.446	.497	2,729,313
Schooling Outcomes			
	Mean	Std. Dev.	Observations
Never Enrolled	.028	.165	2,811,160
Dropped Out	.036	.187	2,811,160
On Track	.813	.390	2,134,088
Drought Exposure			
In-Utero Drought	.177	.382	2,876,063
Drought in Year of Birth	.174	.379	2,876,063

Notes: This table shows summary statistics for the ASER data set as well as exposure to drought from the rainfall data, which we use in subsequent analysis.

Table 2: Effect of In-Utero Drought on Test Scores

Independent Variable: Drought in Utero				
Panel A: Math Scores				
	Coeff	Std. Error	Observations	R-Sq.
Dependent Variable:				
Can Recognize 1 to 9	-.007***	.002	2,186,446	0.7442
Can Recognize 10 to 99	-.026***	.003	2,186,446	0.5704
Can Subtract	-.012**	.003	2,186,446	0.3873
Can Divide	-.002	.003	2,186,446	0.2996
Math Score	-.047***	.006	2186446	0.5476
Panel B: Reading Scores				
	Coeff	Std. Error	Observations	R-Sq.
Dependent Variable:				
Can Read Letter	.004**	.002	2,397,027	0.1725
Can Read Word	-.001	.002	2,397,027	0.3230
Can Read Paragraph	-.007***	.002	2,397,027	0.3751
Can Read Story	-.007***	.002	2,397,027	0.3347
Reading Score	-.009	.006	2397027	0.4343
Panel C: Math Scores, Household Fixed Effects				
	Coeff	Std. Error	Observations	R-Sq.
Dependent Variable:				
Can Recognize 1 to 9	-.007***	.002	2,003,320	0.6974
Can Recognize 10 to 99	-.018***	.002	2,003,320	0.5451
Can Subtract	-.014**	.003	2,003,320	0.4087
Can Divide	-.002	.003	2,003,320	0.3367
Math Score	-.040***	.007	2,003,320	0.5543
Panel D: Reading Scores, Household Fixed Effects				
	Coeff	Std. Error	Observations	R-Sq.
Dependent Variable:				
Can Read Letter	-.002	.002	2,138,351	0.2089
Can Read Word	-.004	.003	2,138,351	0.3589
Can Read Paragraph	-.007**	.003	2,138,351	0.4125
Can Read Story	-.007**	.003	2,138,351	0.3762
Reading Score	-.020**	.008	2,138,351	0.4826

Notes: This table estimates the effect of drought in-utero on test scores. All regressions contain fixed effects for age, year of survey, and are clustered at the district level. Panels A and B contain fixed effects for district, while Panels C and D contain fixed effects for household. Children are marked as having a drought occur while in utero if there was a drought during the monsoon season of the year prior to their birth. Standard errors are reported in parentheses. ***indicates significance at 1% level, ** at 5% level, * at 10% level.

Table 3: Effect of in Utero Drought on Test Scores, by Gender

Independent Variable: Drought in Utero				
Panel A: Math Scores, Boys Only				
	Coeff	Std. Error	Observations	R-Sq.
Dependent Variable:				
Can Recognize 1 to 9	-.008***	.002	1,176,065	0.7549
Can Recognize 10 to 99	-.028***	.003	1,176,065	0.5821
Can Subtract	-.012**	.003	1,176,065	0.3905
Can Divide	-.0001	.003	1176065	0.3064
Math Score	-.048***	.006	1,176,065	0.5546
Panel B: Reading Scores, Boys Only				
	Coeff	Std. Error	Observations	R-Sq.
Dependent Variable:				
Can Read Letter	.005**	.002	1,290,070	0.1730
Can Read Word	-.001	.002	1,290,070	0.3249
Can Read Paragraph	-.007***	.002	1,290,070	0.3768
Can Read Story	-.005**	.002	1,290,070	0.3353
Reading Score	-.008	.006	1,290,070	0.4371
Panel C: Math Scores, Girls Only				
	Coeff	Std. Error	Observations	R-Sq.
Dependent Variable:				
Can Recognize 1 to 9	.004*	.002	983,569	0.7366
Can Recognize 10 to 99	-.001	.003	983,569	0.5610
Can Subtract	-.008***	.003	983,569	0.3769
Can Divide	-.011***	.003	983,569	0.2922
Math Score	-.015***	.006	983,569	0.5429
Panel D: Reading Scores, Girls Only				
	Coeff	Std. Error	Observations	R-Sq.
Dependent Variable:				
Can Read Letter	.002	.002	1,079,822	0.1742
Can Read Word	.0001	.002	1,079,822	0.3218
Can Read Paragraph	-.005***	.003	1,079,822	0.3739
Can Read Story	-.012***	.003	1,079,822	0.3348
Reading Score	-.015**	.006	1,079,822	0.4317

Notes: This table estimates the effect of drought in-utero on test scores by gender. All regressions contain age, year of survey, and district fixed effects and are clustered at the district level. Children are marked as having a drought occur while in utero if there was a drought during the monsoon season of the year prior to their birth. Panels A and B restrict the sample to only male children, and Panels C and D restrict the sample to only female children. Standard errors are reported in parentheses. ***indicates significance at 1% level, ** at 5% level, * at 10% level.

Table 4: Timing of Drought Effects

Independent Variable:	Dependent Variable:	
Age at Drought	Math Score	Reading Score
-2	-.014***	.009
	(.006)	(.006)
-1 (in-utero)	-.047***	-.010
	(.006)	(.006)
0	-.038***	-.001
	(.005)	(.006)
1	-.016***	-.015***
	(.005)	(.005)
2	-.001	.001
	(.005)	(.005)
Observations	2,186,446	2,397,027

Notes: This table estimates the effect of a drought from two years before to two years after birth. All regressions contain age, year of survey, and district fixed effects and are clustered at the district level. Standard errors are reported in parentheses. ***indicates significance at 1% level, ** at 5% level, * at 10% level.

Table 5: Alternative Measures of Rainfall

	Math Score (1)	Math Score (2)
Rainfall (in)	.0001***	
	(.00001)	
Bottom Quintile Rainfall		-.030***
		(.006)
Second Quintile Rainfall		-.014***
		(.005)
Fourth Quintile Rainfall		.015***
		(.005)
Highest Quintile Rainfall		.029***
		(.005)
Observations	2,186,446	2,186,446

Notes: This table measures the effect of two different measures of rainfall on math test scores. The mean of rainfall is 1,286 inches and the standard deviation is 788.5 inches. Both specifications include district, age, and year of survey fixed effects and are clustered at the district level. Standard errors are reported in parentheses. ***indicates significance at 1% level, ** at 5% level, * at 10% level.

Table 6: Effect of In-Utero Drought on Schooling Attainment, by Gender

Independent Variable: Drought in Utero				
Panel A: All Children				
	Coeff	Std. Error	Observations	R-Sq.
Dependent Variable:				
Dropped Out	.0001	.0007	2,451,738	0.0453
Never Enrolled	.002***	.0005	2,451,738	0.0352
On Track	-.020***	.003	1,988,846	0.1354
Panel B: Boys Only				
	Coeff	Std. Error	Observations	R-Sq.
Dependent Variable:				
Dropped Out	.001	.001	1,319,974	0.0426
Never Enrolled	.001	.001	1,319,974	0.0301
On Track	-.018***	.003	1,069,615	0.1286
Panel C: Girls Only				
	Coeff	Std. Error	Observations	R-Sq.
Dependent Variable:				
Dropped Out	-.0002	.0008	1,104,267	0.0499
Never Enrolled	.001	.001	1,104,267	0.0418
On Track	-.008***	.003	898,577	0.1445

Notes: This table estimates the effect of drought in-utero on schooling outcomes. Panel A includes all children, while Panels B and C restrict of boys and girls, respectively. All regressions contain age, year of survey, and district fixed effects and are clustered at the district level. Standard errors are reported in parentheses. ***indicates significance at 1% level, ** at 5% level, * at 10% level.

Table 7: Effect of In-Utero Drought on Child Health Outcomes

Independent Variable: Drought in-utero				
Panel A: Height				
	Coeff	Std. Error	Observations	R-Sq.
Dependent Variable:				
Height (cm)	-7.31**	3.07	18,283	0.39
Panel B: Weight				
	Coeff	Std. Error	Observations	R-Sq.
Dependent Variable:				
Weight (kilos)	-1.03**	.45	18,283	0.35

Notes: All regressions contain fixed effects for year of birth, district, and state. All specifications are clustered at the district level. Children are marked as having a drought occur while in-utero if there was a drought during the monsoon season of the year prior to their birth. The mean of height is 717.1 centimeters and the mean of weight is 81.7 kilograms. Standard errors are reported in parentheses. ***indicates significance at 1% level, ** at 5% level, * at 10% level.

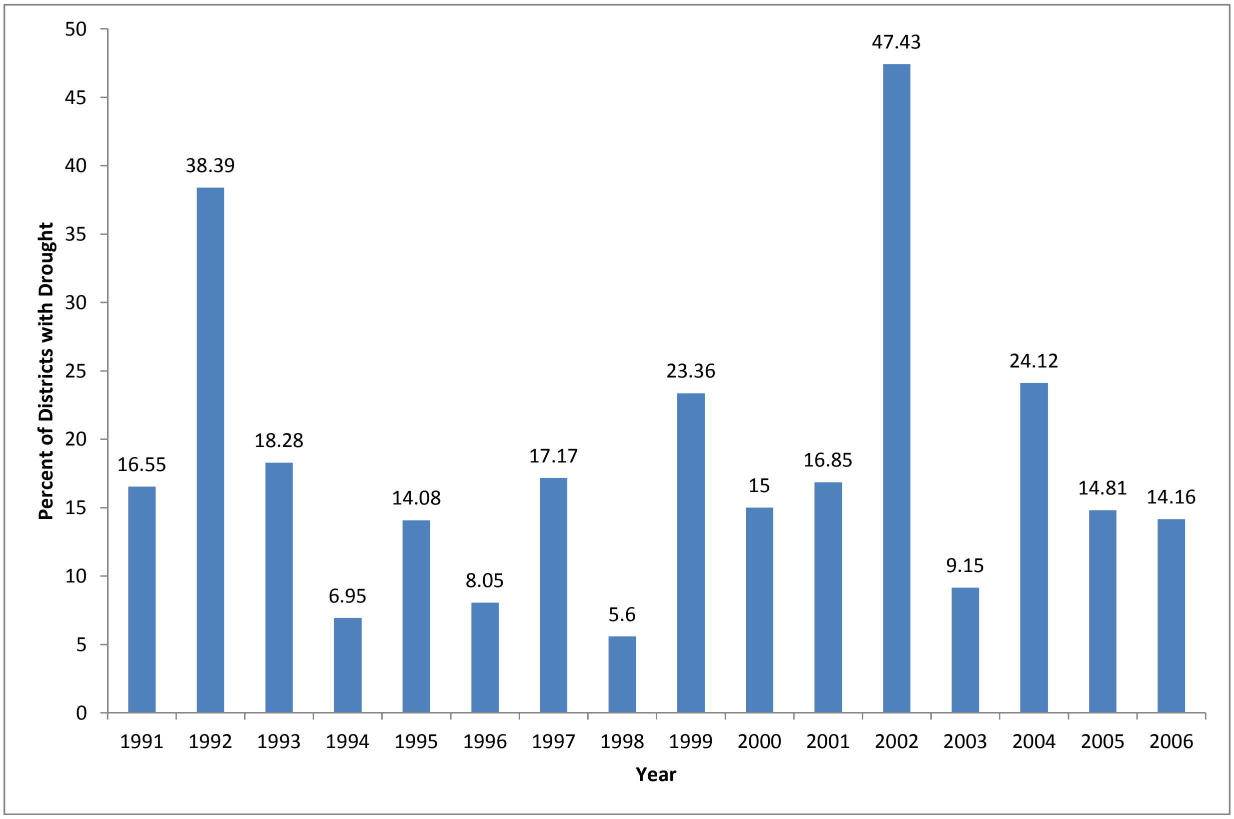

Figure 1: Variation in Drought Across Years

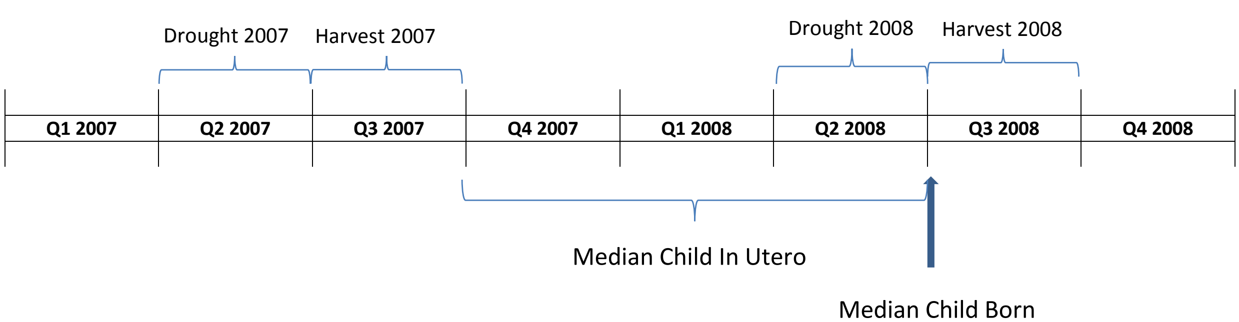

Figure 2: Drought and Gestation Timeline (for 2008)

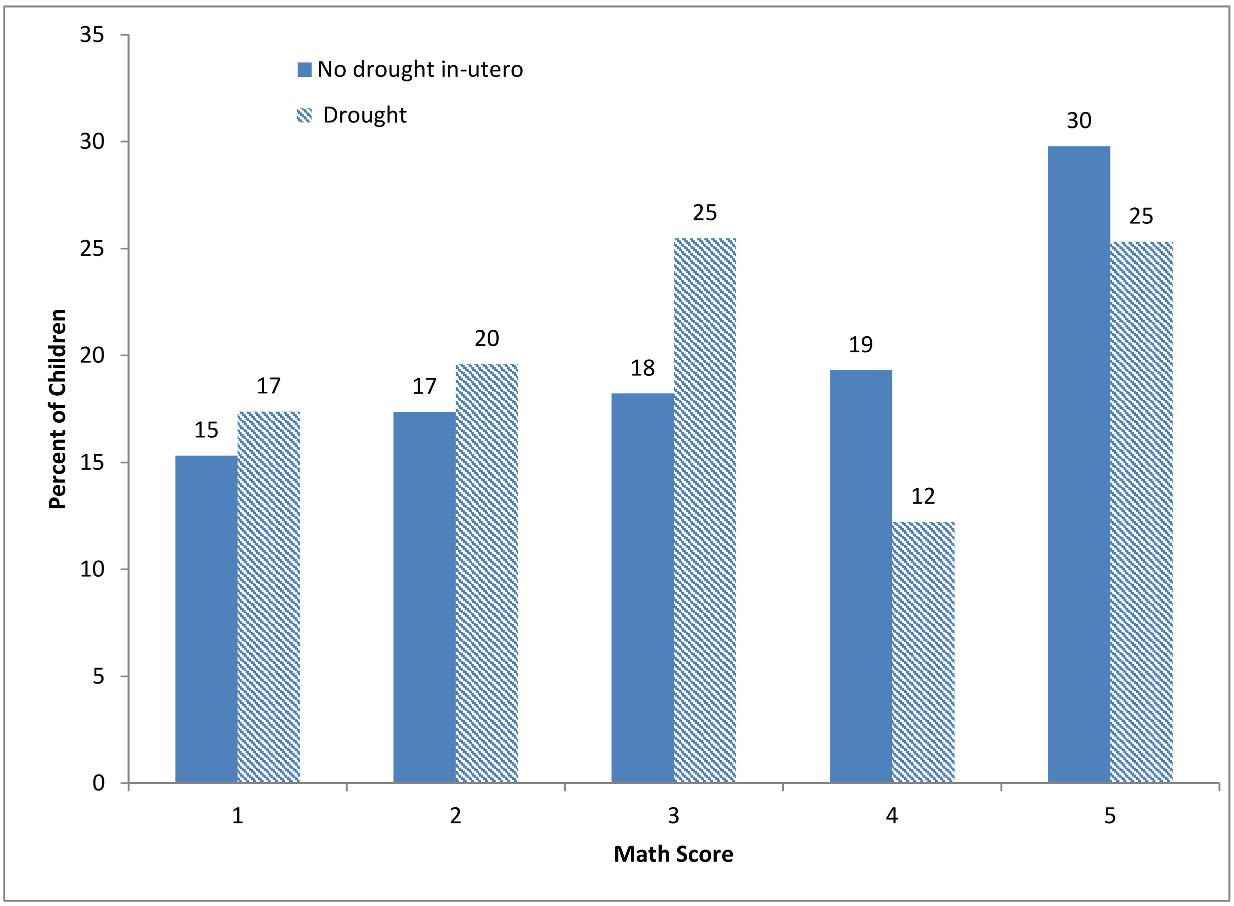

Figure 3: In-Utero Drought Exposure and Math Scores

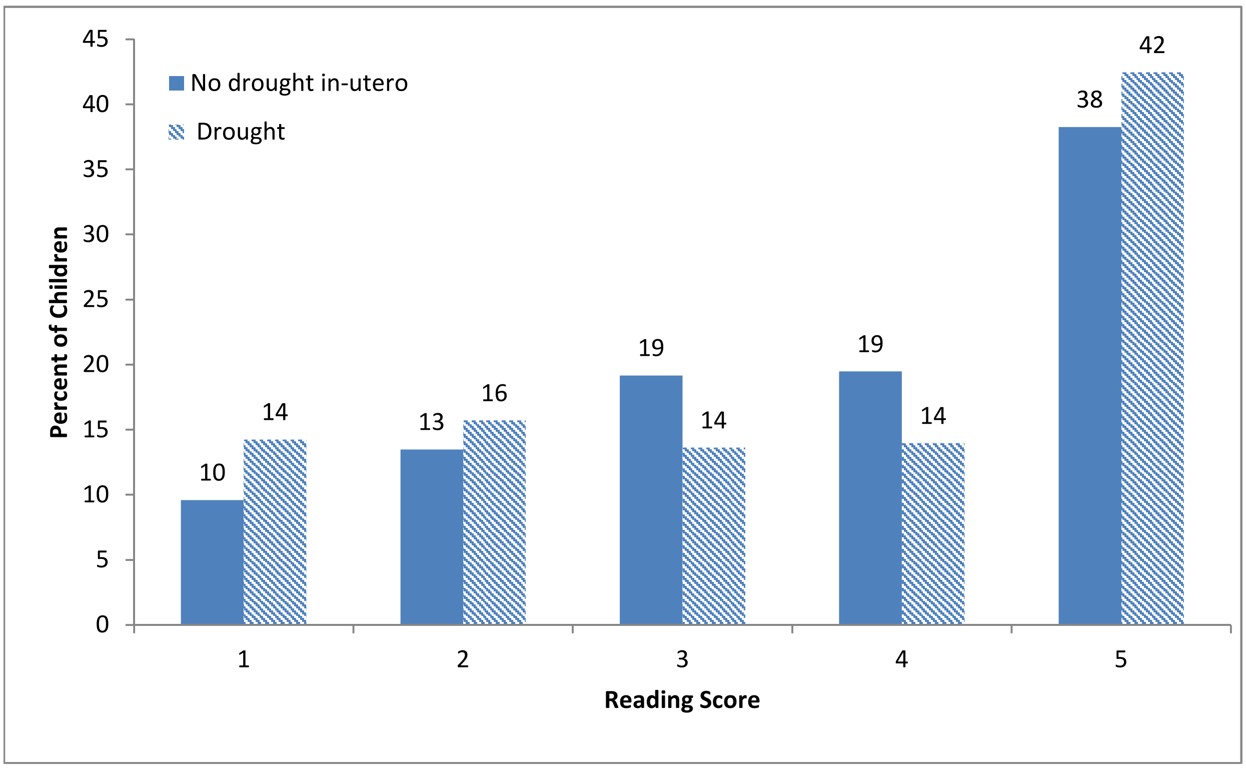

Figure 4: In-Utero Drought Exposure and Reading Scores

A Appendix

Table A1: Testing for Serial Correlation in Rainfall

Dependent Variable: Deviation from district mean this year	(1)	(2)
Deviation from district mean last year	.005	-.031***
	(.011)	(.010)
Year Fixed Effects	NO	YES
Observations	9,248	9,248

Notes: This table tests if there is serial correlation in rainfall in our data. An observation is a district-year. The dependent variable in both regressions is the deviation from mean rainfall in the current year (in inches), where deviation is simply defined as current year rainfall minus the mean rainfall in sample period. The independent variable is deviation from mean rainfall last year (in inches), constructed in the same way. The mean of the deviation is 0 (2.2e-06)and the standard deviation is 223 inches. Standard errors are reported in parentheses. ***indicates significance at 1% level, ** at 5% level, * at 10% level.

Table A2: Does Drought Impact Fertility Decisions?

	(1)	(2)
Drought In-utero (t-1)	.22	.23
	(.15)	(.16)
Drought (t-2)	-.06	-.06
	(.15)	(.15)
Drought (t-3)		-.06
		(.13)
Drought (t-4)		.05
		(.12)
F statistic	8.92	7.26
Observations	2,876,063	2,876,063

Notes: All regressions contain fixed effects for district and standard errors are clustered at the district level. Dependent variable is child's year of birth. Standard errors are reported in parentheses. ***indicates significance at 1% level, ** at 5% level, * at 10% level.