



Network for Studies on Pensions, Aging and Retirement

Netspar DISCUSSION PAPERS

Douglas Almond, Bhashkar Mazumder and Reyn van Ewijk
**Fasting During Pregnancy and Children's
Academic Performance**

DP 12/2011-113

Fasting During Pregnancy and Children's Academic Performance*

Douglas Almond[†] Bhashkar Mazumder[‡] Reyn van Ewijk[§]

Abstract

We consider the effects of daytime fasting by pregnant women during the lunar month of Ramadan on their children's test scores at age seven. Using English register data, we find that scores are .05 to .08 standard deviations lower for Pakistani and Bangladeshi students exposed to Ramadan in early pregnancy. These estimates are downward biased to the extent that Ramadan is not universally observed. We conclude that the effects of prenatal investments on test scores are comparable to conventional educational interventions but are likely to be more cost effective and less subject to "fade out".

*We gratefully acknowledge comments from seminar participants at the Academic Medical Center Amsterdam, ESPE, Erasmus University Rotterdam, London School of Economics, the Nordic Summer Institute in Labor Economics, Northwestern University, the Swedish Institute for Social Research (SOFI), University of Amsterdam, University of Bristol, University of Chicago, University of Mannheim, and "Effects of Early Life Events Conference" at the University of Michigan. We thank Iftikhar Hussain, Francois Kessler, Richard Murphy, Sandra McNally and Felix Weinhardt for helpful comments and the Center for Economic Performance (CEP) at the London School of Economics for providing research facilities. Part of the research was carried out when Van Ewijk worked on the Netspar theme Health and Income, Work and Care Across the Life Cycle. Almond thanks the NSF CAREER program (award no. SES-0847329) for financial support. The views expressed here do not reflect the views of the Federal Reserve system.

[†]Columbia University and NBER

[‡]Federal Reserve Bank of Chicago

[§]University of Mainz, VU University Amsterdam and Tinbergen Institute

1 Introduction

A growing literature has documented that early childhood investments –those made prior to formal schooling years– are more productive in increasing human capital than interventions that come later in life [Heckman and Masterov, 2007]. There is also increasing recognition that investments that occur during the prenatal period may potentially be even more effective than postnatal interventions and significantly less costly to undertake [Doyle et al., 2009]. Targeting interventions during the *in utero* period, therefore, may be particularly compelling on the grounds of economic efficiency.

As emphasized by Heckman and others, learning is a dynamic process that begins well before school: “capabilities beget capabilities” [Heckman, 2007]. Furthermore, human capacity is inherently multidimensional, including health, cognitive, and non-cognitive components which are synergistic over the life course. For example, better health early in life may facilitate learning during school going years. The fetal origins literature has highlighted the staged, developmental nature of early human growth and has emphasized that specific pregnancy sub-periods are thought to imprint distinct physiologic functions. Furthermore, research in developmental neuroscience has demonstrated the greater plasticity of the brain early in life and that there are “sensitive periods” during which particular aspects of cognitive development take place [Doyle et al., 2009]. These brief windows naturally lend themselves to targeted interventions that may be especially cost effective during the prenatal period. This may be contrasted with more conventional educational interventions which may be costly to implement. For example, Project STAR [Krueger, 1999, Krueger and Whitmore, 2001, Chetty et al., 2011], which improved student outcomes through reduced class sizes for a two year period, (on average) cost over \$10,000 per student in 2010 dollars. Finally, a signature feature of fetal-induced changes to health and educational outcomes is their persistence into adulthood. In contrast, the cognitive effects from more conventional education interventions may be subject to “fade out” (e.g. Heckman et al. [2010], Rothstein [2010], Cascio and Staiger [2011]).

Focusing on early antecedents of childhood human capital development, however, is challenging since certain factors are difficult, if not impossible, to modify (e.g. genes). From an empirical point of view, even among the subset of modifiable inputs, it is particularly difficult to distinguish “predictors” from those factors that may actually exert causal effects. Even where causal pathways exist, the eventual effects may remain latent for many years. Therefore, identifying such pathways may require unique data sets that are able to link long-term outcomes to early-life experiences.

In this paper, we argue that the diurnal fast undertaken by pregnant Muslims during the month of Ramadan constitutes a modifiable determinant of prenatal development that has

significant long-term effects on academic outcomes. A central feature of our identification strategy is the fact that Ramadan follows a lunar calendar, and thereby falls on different dates (and seasons) in different years. We present evidence showing that the timing of pregnancy *vis à vis* Ramadan appears exogenous on observable characteristics. We can therefore compare persons born just before Ramadan, and thereby not exposed during pregnancy, to those exposed at different months of gestation using an “intent to treat” approach. As Ramadan lasts for one lunar month, the exposure period is necessarily brief, especially compared to conventional educational interventions.

Our approach departs in four ways from previous design-based observational studies of fetal origins effects. First, we consider an input, the timing of prenatal nutrition, that is relatively manipulable. Although pregnant women are not automatically exempted from fasting, they can request an exception which typically requires them to make up the days later. Many Muslim scholars argue that the Ramadan fast is not obligatory for pregnant women, and, although observance is the norm, fasting rates during pregnancy do vary somewhat across societies. This suggests that there is scope for adaptation in practices.¹ In contrast, other studies have utilized natural experiments, such as famines or influenza pandemics, where variation in the environment is “caused by conditions outside the control of the mother” [Currie, 2009]. While extreme natural events provide credible sources for identification, they are less relevant when considering policies that may modify individual behavior. Second, since the Ramadan fast primarily affects the diurnal timing of nutrition, it constitutes a far less extreme treatment than the famine episodes, pandemics, and natural disasters previously analyzed. Therefore, it is informative about whether less extreme shocks to the fetal environment also have long-term effects. In particular, meal skipping, “morning sickness” and dieting during pregnancy (especially prior to pregnancy recognition) are fairly common in developed countries and are comparable to fasting in terms of effects on the timing of nutrition and their effects on the intrauterine environment. Thus, the identified linkages to the prenatal period may generalize to other populations. Third, as most Muslims were *in utero* during a Ramadan, and fasting during pregnancy remains common today, the population affected by fasting is much larger than those afflicted by historical famine episodes or disease outbreaks, i.e. the estimates of fasting’s impact is of interest *per se*. Finally, in contrast to previous studies which have examined long-term adult outcomes, we focus on measures of human capital at age 7 to shed light on how fetal-induced effects are first manifested during the primary school years. In this respect, our analysis is novel compared with other recent studies of the long-term effects of Ramadan fasting (Almond and Mazumder [2011], Van Ewijk [2011]).

¹For example, at Sorrento Maternity Hospital in Birmingham, England, three quarters of Muslim mothers reported fasting during pregnancy that coincided with Ramadan [Eaton and Wharton, 1982].

Our study uses school register data from England containing the national “Key Stage 1” assessments in math, reading and writing. We use Pakistani or Bangladeshi ancestry to identify Muslim students and estimate difference in difference models to estimate the effects of Ramadan exposure on Muslims compared to non-Muslims during the *in utero* period. Our main finding is that Muslim students exposed to Ramadan fasting in the first trimester of their mothers’ pregnancy have significantly lower achievement scores. For example, we find that the math scores of students who were exposed to Ramadan during the first trimester are reduced by 0.06 to 0.08 standard deviations. To the extent that not all pregnant Muslim women are fasting, our estimates understate the true effect size and serve as a lower bound. The magnitudes of the effects are comparable to many conventional educational interventions such as the effects of charter schools, Teach for America or Head Start [Dobbie and Fryer, 2011]. Relative to these interventions, however, the potential costs of altering behavior may be significantly lower.

2 Literature Review

2.1 Economics

From an economic perspective, one can gain insight into the potential importance of early life experiences through a model of human capital production. As suggested by Heckman, the different stages of childhood can be viewed as distinct inputs into the production of subsequent “capacity”. Let I_1 denote investments occurring during the prenatal period and I_2 investments during the postnatal period (i.e. the rest of childhood). Given that cognitive development early in life occurs in distinct stages, I_1 and I_2 are likely imperfect substitutes in the production of capacity. In the extreme case of a Leontieff technology, human capacity cannot exceed that determined by the minimum of investments during the prenatal period [Heckman, 2007].

Furthermore, if there are multiple dimensions to “capacity” (e.g. cognitive, non-cognitive, health) then there can be synergies across these dimensions. For example, a child born in better health may have an advantage in creating cognitive and non-cognitive capacity. Finally, the production technology may incorporate dynamic complementarities [Heckman, 2007] whereby investments in stage t of childhood are more productive when there is a high level of capability in stage $t-1$.

Economists have also sought to establish links between prenatal conditions and human capital outcomes empirically. For example, Currie and Hyson [1999] using the British National Child Development Survey found that the pass rate for math and English O-level tests was roughly 25% lower for low birth weight children. More recent design-based empirical studies by economists have found that human capital outcomes respond to a range of

prenatal shocks, particularly those experienced during first half of pregnancy. Field et al. [2009] evaluated the effect of prenatal iodine supplementation on subsequent educational attainment in Tanzania, finding that supplementation during the first trimester increased completed schooling by as much as half a year. These effects persisted in a siblings comparison, and were generally stronger among girls. Almond et al. [2009] studied prenatal exposure to radioactive fallout from the 1986 Chernobyl meltdown on middle school performance in Sweden. Exposure to ionizing radiation between weeks 8 and 25 of gestation reduced the likelihood of qualifying for high school by 3% and reduced math grades by 6% (the measure closest to IQ). Interestingly, no health effects of radiation exposure were detected in the prenatally exposed cohorts, suggesting the damage was “subclinical.” Kelly [2001] considered the impact of prenatal exposure to the fall 1957 “avian flu” pandemic in Britain, using the serendipitous timing of the 1958 British cohort study (born March 1958). Kelly [2001] found negative impacts on test scores that interestingly appeared independent from the negative impact of the 1957 pandemic on birth weight.

While recent studies have successfully exploited natural experiments to demonstrate causal pathways, the incidences of such prenatal experiences are relatively rare and are not likely to be the main determinants of current modifiable differences in the prenatal environment, particularly in developed countries. As we describe in the next section, the biophysical changes induced by Ramadan fasting during pregnancy resemble those occasioned by other determinants of nutrition timing in developed countries such as meal skipping, dieting, and nausea and vomiting (“morning sickness”).

2.2 Biological Mechanisms

Almond and Mazumder [2011] and Van Ewijk [2011] provide overviews of the biomedical literature concerning the potential pathways between prenatal fasting and long-term outcomes. We briefly review some of the mechanisms that may be particularly relevant for cognitive function. One potential pathway arises through a set of biochemical changes known as “accelerated starvation” that occurs in pregnant women who undergo an extended period of fasting. Pregnant women experience pronounced declines in blood glucose levels and sharp increases in ketones and free fatty acids as they begin to metabolize their stores of fat. Such conditions can arise in as little as 12 hours and studies have documented these changes during the Ramadan fast in both developed and developing countries (Prentice et al. 1983, Malhotra et al. 1989). Animal studies have linked exposure to ketones early in pregnancy to neurological impairments (e.g. Hunter and Sadler 1987) and studies of humans have associated ketone exposure in diabetic mothers to diminished cognitive ability [Rizzo et al., 1991].

The literature on the developmental origins of health and adult disease has emphasized

how environmental exposures in pregnancy, such as nutritional disruptions, can lead to permanent alterations in the body's systems in order to improve the likelihood of survival to reproductive age in the perceived environment at birth [Gluckman and Hanson, 2005]. These "predictive adaptive responses" (PARs) make individuals more prone to poor health in adulthood. The most well-known examples relate to heart disease and diabetes, but there may be other manifestations of PARs as well. Although the literature has only begun to speculate at the precise mechanisms behind PARs, disruptions to the flow of glucose are thought to be one of the key signals of poor environmental condition during fetal development. Gluckman and Hanson note that "the developing embryo will change the relative assignment of cells to the inner cell and outer cell mass according to whether it perceives a problem in glucose supply" [Gluckman and Hanson, 2005, p31-32].

One particular example of a PAR that has received significant attention in the literature is the notion that prenatal nutritional deprivation or maternal stress can lead to alterations in the neuro-endocrine system or "HPA axis" which in turn, can lead to permanent health effects. PARs operating through the HPA axis are notable for our purposes for two reasons. First, several studies have linked maternal stress during pregnancy to behavioral and cognitive deficits in children [Kapoor et al., 2006, LeWinn et al., 2009, Aizer et al., 2009], and it is hypothesized that this may be due to modifications to the HPA axis. Direct evidence linking the HPA axis to cognitive impairments has been found in animal studies. Second, a recent study documented elevated levels of the hormone cortisol, which occurs when there is heightened sensitivity in the HPA axis, among pregnant women who fasted during Ramadan [Dikensoy et al., 2009]. It is also worth noting that another recent study co-authored by David Barker, one of the pioneering epidemiologists in the fetal origins field, linked Ramadan fasting to alterations in placental growth due to fetal programming [Alwasel et al., 2010].

The brain is thought to be especially susceptible to the fetal environment due to the complexity of its development. Gluckman and Hanson write: "this complexity means that the fetal brain is very sensitive to environmental stimuli that might irreversibly damage it" [Gluckman and Hanson, 2005, p46]. They further note that "the number of neurons is almost entirely determined in fetal life and is largely completed in mid-gestation" [Gluckman and Hanson, 2005, p46] suggesting that nutritional shocks in the first half of pregnancy may be especially harmful. They further point out that the fetal environment may play a contributing role in the development of certain psychiatric diseases.

2.3 Long-term Effects of Ramadan Fasting

We are aware of three studies that examine effects of prenatal exposure to Ramadan fasting on childhood or adult outcomes.² Azizi et al. [2004] found no statistically significant effects of fasting on the IQ scores of 191 children between the ages of 4 and 13 attending 15 primary schools in Iran. The study compared 98 treated children whose mothers fasted for at least 27 days during Ramadan, with 93 control children whose mothers did not fast at all during Ramadan. However, mean differences between the treatment and control groups were found in certain characteristics, such as breast feeding duration and socioeconomic status, that were either statistically significant or quantitatively meaningful.³ In addition to having a relatively small sample, the study appeared to have selected cases based on potential outcomes which could have imparted some bias.⁴

More recently, Almond and Mazumder [2011] linked Ramadan fasting to adult outcomes in Uganda and Iraq using Census data. They find that full exposure to Ramadan fasting in the first month of pregnancy increased the likelihood of a disability by about 20% with especially large effects on mental/learning disabilities. Van Ewijk [2011], also focusing on adult samples of Indonesians from the Indonesian Family Life Survey (IFLS), found effects of fasting on a variety of measures of health including coronary heart problems and diabetes. Both Almond and Mazumder [2011] and Van Ewijk [2011] utilize a research design that compares Muslims whose *in utero* period overlapped with Ramadan to Muslims who were unexposed and show that pre-determined observable characteristics do not vary with exposure.⁵

²There are a number of studies in the biomedical literature that examine the effects of fasting during Ramadan on fetal and birth outcomes which are discussed in Almond and Mazumder [2011]. As noted here, most previous studies rely on the strong assumption that non-fasters are comparable to fasters at a point in time and typically use samples that lack sufficient power to detect small but quantitatively meaningful effects. In contrast, using the universe of natality data on 18 birth cohorts from the U.S. state of Michigan, Almond and Mazumder [2011] find significant effects of fasting on lowering birth weight and reducing the likelihood of a male birth.

³The duration of breastfeeding was about three months longer in the treated group (statistically significant at the 5 percent level). An index of socioeconomic status as well as income and home ownership were also all higher in the treated group, though not statistically significant. For example, 18 percent of the treated owned their own home compared to 13 percent of the control group.

⁴Of the 141 children who could have been included in the treated group the sample of 98 included *all* of those who fasted in the third trimester but only a sample of those who fasted earlier in pregnancy. The oversampling of those with late exposure is problematic since the neuro-development literature has emphasized the importance of early exposure. Further, they appear to have selected sample members such that mothers “with any history of problems such as drug consumption, smoking, and thyroid dysfunction during pregnancy, dystocia, and other problems during different stages of development affecting children’s IQ from the fetal stage to childhood were excluded”.

⁵Van Ewijk [2011] also finds that the results are robust to including mother fixed effects suggesting that any unobservable forms of selection would have to be sibling-specific.

3 Data and methodology

3.1 Data

In England, all students attending state schools are assessed at different points in their schooling career, or “Key Stages”, to measure their academic performance in different subject areas. We start with the population of students who were assessed at Key Stage 1 between 1998 and 2007 when they were approximately 7 years old. The Key Stage 1 score is based on a teacher assessment of the students’ proficiency in reading, writing and mathematics.⁶ Teacher assessments are made following detailed guidelines based on “National curriculum levels” that describe levels of proficiency in each subject area. The assessment is based on a combination of tests and tasks that take less than three hours to administer.⁷ Students at Key Stage 1 should be at level 2. The teacher assessment can take on one of the following values: 1, 2C, 2B, 2A, 3, 4 or above.⁸ Following Department of Education guidelines, we translate these assessments into numerical scores which we then transform in z -scores using the full sample.^{9,10} We also use as an aggregate measure the first principal component from a principal components analysis (PCA) on math, reading and writing. This variable captures 83% of the total variance of the three constituent subjects.

We use a unique student identifier to link the Key Stage 1 scores to other student level data contained in the Pupil Level Annual School Census (PLASC). The PLASC is constructed based on electronic records provided by each school in England to the Department for Education and Skills (DfES) and covers all enrolled pupils as of January of each year. Starting with the 2002 data we link the Key Stage 1 scores to the PLASC for that year. However, prior to 2002, some of the key background characteristics we require such as ethnicity are unavailable in the contemporaneous PLASC. Instead we link these individuals

⁶Prior to 2004, our data contained both teacher assessments *and* standardized tests (“National Curriculum tests”) but starting in 2005 we only have the teacher assessment. Therefore in order to have a uniform measure across all years we use the teacher assessment measure. Since students at this level have only 1 teacher and since the teacher gives these assessments at the end of the school year, when he or she knows the child well, this may actually provide a more reliable measure than standardized test scores which contain considerable noise [Kane and Staiger, 2002].

⁷Schools and local areas have some discretion as to how the teacher assessments are conducted. The following link provides a description of the assessment process: http://www.direct.gov.uk/en/Parents/Schoolslearninganddevelopment/ExamsTestsAndTheCurriculum/DG_10013041

⁸Level 4 for math and reading only exists from 2002 on. Levels 2A, 2B, and 2C do not exist before 2004 (only level 2 exists in these years). In addition, students can be assessed as: “W” for a child who is working towards level 1; “A” for a child not assessed due to absence, or a child who has had a long period of absence, or there is insufficient information to enable a teacher assessment result to be calculated; or “D” for a child for whom teacher assessment has been “disapplied”.

⁹W = 3 points; level 1 = 9 points; level 2C = 13 points; level 2B = 15 points; level 2 = 15 points (where no breakdown of level 2 reported); level 2A = 17 points; level 3 = 21 points; level 4 = 27 points, see <http://www.dcsf.gov.uk/rsgateway/DB/SFR/s000867/SFR21-2009.pdf>

¹⁰Standardizing the scores per cohort instead of over the full sample gives virtually the same results as those presented below.

through a 2-step process to the 2002 PLASC to obtain their characteristics as of 2002.¹¹

3.2 Difference in difference strategy

Previous work on adult outcomes [Almond and Mazumder, 2011, Van Ewijk, 2011] has used an “intent to treat” (ITT) design that compares the outcomes of Muslims who were *in utero* during Ramadan to those who weren’t. One important issue in this identification strategy is separating seasonal factors from true fasting effects since it is well established that season of birth (or conception) has long-run effects [Doblhammer and Vaupel, 2001]. This is particularly important in the context of studying educational outcomes because the age cutoffs that determine school entry lead to a sharp discontinuity in school performance by timing of birth. This is apparent in England as can be seen in Figure 1 where those born on or after September 1 have significantly higher Key Stage 1 math scores.¹² It is also clear that in addition to this discontinuity, there is a notable trend reflecting the age at which one is assessed.

The prior studies took advantage of the gradual movement over time of Ramadan throughout the Gregorian calendar year. Since Ramadan is based on the lunar calendar, it begins about 11 days earlier each year. Therefore, with data covering three decades of birth cohorts, one can employ standard seasonal controls (e.g. month dummies) and readily remove confounding effects since Ramadan would have made a complete cycle over the year. In these prior studies the effects are identified by comparing Muslims whose prenatal period overlapped with Ramadan to those who were never *in utero* during Ramadan, and the same estimation approach can be applied separately to non-Muslims as a falsification check.

In our case, however, we have only ten birth cohorts whose potential prenatal Ramadan exposure only could have occurred during one of five winter and spring months (December to April) and so it is not obvious that seasonal controls are sufficient. Given the limited number of cohorts and the strong seasonal effects for educational outcomes, we instead utilize a “difference in difference” strategy where we take the effect on Muslims and further subtract any effects for non-Muslims that may arise due to possible seasonal effects. The number of birth cohorts also motivates our use of the Key Stage 1 assessments since even fewer cohorts would be available if we used Key Stage 2 test scores in our sample. As we show later,

¹¹We first must link these students to the 2005 PLASC where we are able to retrieve an identifier that allows us to link them back to 2002. This imposes a requirement that the students who take the Key Stage 1 prior to 2002 must have remained in the English school system through 2005. We do not think that this selection rule is much of a concern since the students who took the test as early as 1998 would still only be 14 years old as of 2005 and therefore highly likely to have remained in school. We will however, remove students who either left England or left the state school system by 2005.

¹²For example, those born on August 31, 1999 would have received the Key Stage 1 assessment in 2006 whereas those born on September 1st 1999 would receive the assessment in 2007. Note that the September 1st cut-off is strictly observed, and that retention and grade skipping are very rare at this age: only 0.28% of all students are not in their expected cohort.

it turns out that the differencing is not critical. There are no systematic Ramadan effects for non-Muslims, echoing the “placebo” results from the prior studies. As we will show the effects remain even if we only use Muslims who were not exposed to Ramadan as the only control group as in the previous studies.

Ideally, for our treatment group we would like to identify Muslim students who were *in utero* during Ramadan. Since the PLASC does not identify the religion of the student, we instead assign Muslim status to students who report their ethnicity as Pakistani/Bangladeshi. According to the 2001 Census, 92 percent of Pakistanis and Bangladeshis report that they are Muslims. To reduce the scope for measurement error, we use only those Pakistani/Bangladeshis who are living in a region (local authority) where at least 90 percent of Pakistani/Bangladeshis with a reported ethnicity are Muslims according to the Census.¹³ We suspect that there is still some residual measurement error since the reported ethnicity of students is not always constant across years. Figure 2 shows the distribution of Pakistani and Bangladeshi in England. Areas with high concentrations of Muslims include London and the areas around Birmingham in the West and around Manchester in the North-West.

For our control group we use Caribbean students since they are a comparable ethnic minority group. Caribbeans have similar levels of school performance and nearly identical rates of free school meal status –a proxy for socioeconomic status. In Table 1 we show that the average scores of Caribbeans are 0.20 standard deviations below the national average compared to -0.36 for our designated Muslim students. About 35% of both groups of students receive free school meals. This compares to an average rate of free school meals of about 16% for white British students. Although Indians are culturally more similar to Pakistanis and Bangladeshis, a sizable minority are Muslim and hence they would be a “contaminated” control group. Further, Table 1 shows Indian students outperform the national average and are less likely to receive free school meals.

An issue that arises in classifying Caribbeans is that the ethnicity codes were expanded in 2003 to create a separate category for mixed-race Caribbeans (“white and black Caribbean”) in addition to the traditional category of “Caribbean”. For our main analysis we have combined both groups in order to maximize our sample. This leads to a large increase in the number of Caribbeans starting in 2003. As a robustness check we have also excluded mixed race Caribbeans and find similar, though less precise, results. As we discuss later, we have also run all of our models using white British students as an alternative control group and find similar, and much more precisely-estimated effects.

¹³This removes only about 1.2% of all Pakistani/Bangladeshis. We also drop any students who report a mixed ethnicity of White and Bangladeshi or White and Pakistani.

3.3 Ramadan Measures

We collected the start dates and end dates for the relevant Ramadans affecting our birth cohorts. In order to identify whether Ramadan overlapped with the *in utero* period, we use one’s exact birth date and assume a normal gestation length of 266 days (since conception) for each individual. We then create a set of indicator variables to identify when during gestation Ramadan began. For each of the nine months of pregnancy we generate a separate variable (e.g. Month 1, Month 2, ..., Month 9) to indicate whether Ramadan began during that month of pregnancy. In addition, we create a “Month 0” variable to capture conception during Ramadan and early gestation exposure to fasting. Those whose pregnancies, by this calculation, do not appear to overlap with Ramadan are further subdivided into two categories. We classify individuals as “probably not exposed” if they were conceived within 14 days after Ramadan had ended and “certainly not exposed” if they were conceived more than 14 days after Ramadan. In Figure 3 we provide an example of how various pregnancies would be classified based on the exact date of birth and the timing of Ramadan. Our reliance on the normal gestation length creates some potential measurement error for most of these indicators of Ramadan exposure since some pregnancies will be preterm or longer than full term.¹⁴

The size of our coefficients can be interpreted as the magnitude of fasting’s effect only if *all* Muslims who were pregnant during a Ramadan chose to fast. Since fasting rates typically depart from unity, our ITT approach *underestimates* the treatment effect of fasting and can be viewed as a lower bound. Our estimates can be rescaled by multiplying by the inverse of the fasting rate in order to approximate the treatment effect of fasting.

3.4 Specification

Our main specification uses ordinary least squares to run regressions of the Key Stage 1 assessments in math, reading and writing, as well as their first principal component, on the Ramadan exposure measures. Since the three subject tests are *z*-scores, the coefficients can be interpreted as the effect sizes in standard deviation units. However, the standard deviation of the first principal component is 1.58, so the estimated effects for this outcome should be scaled down by this factor in order to be comparable to the other coefficients. The excluded category is those who are classified as “certainly not exposed”, so all effects are relative to this group. Additional controls include month of birth dummies, a dummy

¹⁴Our Month 9 variable captures individuals who were born during Ramadan. To the extent that the exact date of birth is measured accurately there should be no misclassification. Similarly, those identified as “certainly not exposed” would only be misclassified if the term of gestation exceeded 280 days which is rare [Kieler et al., 1995]. In that case conception would overlap with the end of Ramadan. We note that as long as the date of birth is not incorrect, premature births will never be misclassified as “not exposed” if they actually were exposed but it is possible that they could be misclassified as “exposed” even if they weren’t.

for female, a dummy for Muslims, a dummy for free school meal eligibility and a set of geographic dummies for each “Census output area”.¹⁵ To further address concerns about time trends we also include a cubic in the number of days between the date of birth and January 1 1960. Standard errors are clustered at the school level. We fully interact each regressor (except for the geographic fixed effects) with a dummy for “Muslim”. This allows for example, for separate time trends and seasonal patterns for Muslims and Caribbeans, (albeit with possibly imperfect seasonal control). The coefficients on the Muslim interaction terms for the Ramadan measures are the main objects of interest. For example, a significant negative coefficient on the interaction between being Muslim and being *in utero* during the first month of gestation would suggest an effect of fasting that would be over and above any effect that might be found for Caribbeans due to confounding factors such as residual seasonality that may remain after the interacted calendar month-of-birth fixed effects.

We also estimate the same specification using free school meal eligibility as the dependent variable to show that there is no selective timing of pregnancies related to socioeconomic status. Finally, as a “placebo treatment”, we estimate any prenatal Ramadan “effects” for Caribbeans using white British as a control group, where we should not expect to see any effects. This helps ensure that our results are not driven by any other misspecification of the model such as any residual seasonal effects that might be correlated with the timing of Ramadan.

4 Results

4.1 Difference-in-differences vs. Caribbeans

We begin by showing our main results in Table 2 where we present the coefficients on the Ramadan measures interacted with an indicator for being Muslim. These show the effects on Muslims of Ramadan starting in each month of pregnancy compared to Muslims with no *in utero* exposure, relative to Caribbeans. The results show consistently significant negative effects of exposure in the first three months of pregnancy. For example, column (1) shows that Muslim students exposed to start of Ramadan in the first month of pregnancy have Key Stage 1 assessments in math that are 0.068 standard deviations lower. Similarly sized effects are found in the month of conception and the second and third months of pregnancy and for the same four periods for reading assessments (column 2) and writing assessments (column 3). The implied effect sizes for the first principal component of these subjects from principal components analysis (PCA) shown in column (4), when converted into standard deviation units, is also very similar. The largest effects appear to be in the third month of

¹⁵Our sample includes 51,187 census output areas. Output Areas are the base unit for the release of Census data and are based on common demographic, household and economic characteristics.

pregnancy when the effects on math and on the first principal component are about 0.08 standard deviations. This accounts for about 20% of the overall test score gap between Muslims and the national average in these subjects. As we discussed earlier, the fact that not all pregnant Muslim women observe the fast implies that our estimates should be viewed as a lower bound. The similarities in the pattern of effects across the outcomes is illustrated in Figure 4 which plots the coefficients shown in columns 1 through 4 of Table 2.¹⁶

Column (5) shows the results on math scores using a specification in which the geographic fixed effects have been omitted. Comparing columns (1) and (5), we find the results are actually a bit stronger once we control for location fixed effects, suggesting that our results aren't driven by geographic differences. In column (6) we only use Muslim sub-sample and no longer estimate a difference in difference model. In this specification our identification is based on only the effects of Ramadan exposure relative to other Muslims whose *in utero* period did not overlap with Ramadan. We again find that months 1 through 3 all show negative effects that are significant at the 5 percent level and that those conceived during Ramadan have scores that are significantly lower at the 10 percent level.

Table 3 shows the coefficients on Caribbean students, our control group. Importantly, we find no instances of negative effects of Ramadan exposure on our various outcomes that are quantitatively or statistically meaningful.¹⁷ This is reassuring since it suggests that the negative effects on Muslims are not driven by other factors such as residual seasonality that may be confounded with Ramadan exposure.

4.2 Evidence on Selective Timing of Pregnancies

A crucial assumption for our identification strategy is that there is no systematic selection with respect to the characteristics of Muslims who conceive relative to the timing of Ramadan. For example, if there were some reason that Muslims of lower socioeconomic status were more likely to conceive in the three months prior to Ramadan, then this might provide an alternate explanation for our findings of strong effects in the first trimester. A detailed analysis of selection on observables by Almond and Mazumder [2011] using Michigan natality data found no evidence of selection bias in terms of the timing of pregnancies relative to Ramadan using variables such as parental education, maternal smoking behavior or a

¹⁶We also find similar results if we use the original categorical coding of the assessment levels using ordered probit models or if we run linear probability models with an indicator for attainment of level 2 as the outcome.

¹⁷We do find that 3 out of the 44 coefficients in columns 1 to 4 of Table 3 show effects that are statistically significant at the 5 percent level which is roughly what one might expect to find purely by chance. However, all of these are positive effects and none occur in months 1 through 3 where we find our largest effects on Muslims. For Muslims, 13 of the 44 coefficients in Table 2 are negative and significant at the 5 percent level and another 7 are negative and significant at the 10 percent level. Importantly every coefficient for Muslims in the first trimester (months 0 to 3) is negative and statistically significant at either the 5 or 10 percent level.

Medicaid receipt (a proxy for income). Van Ewijk [2011] finds no differences in parental health, income, and assets between Indonesian children by their *in utero* exposure during a Ramadan. Van Ewijk [2011] further shows that estimates of health effects of prenatal Ramadan exposure on children’s health are robust to the inclusion of mother fixed effects, suggesting that any forms of selective timing of pregnancy must have been specific to each child.

We re-examine the possibility of selective timing of pregnancy with our British educational data by running our statistical model using Free School Meal status (FSM) as a dependent variable. Free School Meal eligibility is a commonly used proxy for low socioeconomic status (SES) since it is means tested. The results of this exercise are shown in column (7) of Table 2. We find that Muslims who were *in utero* during the first three months of pregnancy when Ramadan began the period where we find consistent effects on achievement were no more likely to be eligible for FSM. Further looking at months 1 through 9, five of the months actually have negative coefficients suggesting that Ramadan was associated with lower rates of FSM, or higher SES.

We do note, however, that Muslims exposed to Ramadan in the seventh month of pregnancy and those who were conceived during Ramadan are slightly more likely to receive FSM. While it is possible that the latter result could reflect some type of actual behavioral difference during Ramadan between low and high socioeconomic status Muslims (perhaps because of differences in levels of observance or differences in sexual practices) we are somewhat doubtful of this. One reason for our skepticism is that those who we label as probably not exposed, many of whom were conceived well after Ramadan ended, have even *higher* rates of FSM. It would be surprising if lower socioeconomic status Muslims were more likely to conceive both during Ramadan and after Ramadan ended, if this was due to a behavioral difference associated with Ramadan. Finally, even if it were the case that more low income Muslim women systematically chose to conceive during Ramadan, this would not explain the pattern of results we find of negative effects for those women for whom Ramadan began during one of the first three months of pregnancy, and whose conception therefore, preceded Ramadan.

4.3 Robustness Check Using White British

To benchmark our difference-in-difference estimates using an alternative control group, we have also run the same specification using white British students rather than Caribbeans. Although this comes at the expense of using an arguably less comparable control group, it increases the sample size to around 4.6 million observations and thereby provides much greater precision. The results for the three subject tests as well as for the principal component are plotted in Figure 5. We once again find that early exposure is associated with lower

assessments in all subjects. For example, for those Muslim students for whom Ramadan began in the third month of gestation, math scores are lower by about 0.06 standard deviations which is similar though somewhat smaller in magnitude than the 0.08 standard deviation effect when using Caribbeans as the control group. For the first principal component, the implied effect size is now about -0.054 in standard deviation units, compared to -0.076 as implied by the results in Table 2.

4.4 Robustness Check: Placebo Treatment

As an additional check on our identification strategy, we conducted a “placebo” test by estimating the same difference in difference regression using Caribbeans as the treatment group and white British as the control group. We would be concerned if we found “effects” on the Caribbean group who we know are not observing Ramadan. Any such placebo effects might suggest that our specification is not adequately dealing with seasonality. The results are shown in Figure 6. We find that only 1 out of the 44 coefficients is statistically significant at the 5 percent level. Further, there does not appear to be any systematic pattern in the results that would suggest that our first trimester effects on Muslims are in any way an artifact of seasonality.

5 Conclusion

The Ramadan fast lasts just one lunar month, yet our results suggest that alterations to the timing of prenatal nutrition during this brief period potentially has ramifications on lifelong human capital. Most commonly studied childhood determinants of human capital reflect investments that occur over much longer periods of time, may be subject to “fade out”, and are much more costly to undertake. Further, the magnitude of the effects of prenatal exposure to Ramadan are very similar to the size of the treatment effects of “later-life” interventions. Dobbie and Fryer [2011] summarize the effects of successful educational interventions and the magnitude of their effects on student performance. Charter schools in New York were found to increase test scores by .09 standard deviations, Teach for America raised math and reading scores by 0.15 and 0.03 standard deviations (respectively), and Head Start increased scores on applied problems by .15 standard deviations. Our lower bound estimates suggest that fasting in the first trimester of pregnancy reduces academic scores by between 0.05 and 0.08 standard deviations. Despite its brevity, exposure to Ramadan during an especially sensitive developmental period exerts meaningful and persistent effects on human capital accumulation. This suggests relatively low cost investments in prenatal nutrition may yield high returns.

We find that the effects are statistically significant and largest in the first three gestation

months. While this is consistent with the hypothesized predictions of much of the developmental origins literature, it is possible that some of the gradient in the effect size may be due to differential rates of fasting during the course of pregnancy, for which we unfortunately have no data. Presumably fasting observance would be highest shortly after conception when many women do not yet know they are pregnant. Therefore, our ITT estimates of first month exposure are probably closest to fasting's effect. Interestingly, however, our effects appear to rise monotonically over the course of the first trimester. Indeed, the largest effects appear to be in the third month of gestation when we speculate that Ramadan would not be universally observed, suggesting that the effects on cognitive development may be particularly large at this stage. Future research that can combine information on fasting behavior over the course of pregnancy with a credible research design may be able to better sharpen our understanding of these patterns of effects.

The only previous study to consider the effects of prenatal Ramadan fasting on *school age* outcomes found no effect [Azizi et al., 2004].¹⁸ If as we suspect, the human capital effects we find are unknown to Muslim parents, postponing the Ramadan fast until after pregnancy may offer a low cost route to improved outcomes. The fact that Ramadan fasting alters the biochemical characteristics of the intra-uterine environment in a way similar to other restrictions on the timing of prenatal nutrition, suggests that these effects may also generalize to non-Muslims. Future research should seek design-based approaches to assess the effects of dieting, meal-skipping (reported by 24% of pregnant mothers in US, see Siega-Riz et al. [2001]), and nausea and vomiting during early pregnancy, which can all affect the timing of nutritional intake. In the case of dieting and meal skipping, these potentially harmful behaviors tend to be more common early in pregnancy, especially prior to the pregnancy being recognized, see e.g. Ebrahim et al. [2000]. In addition, future research might consider the effect of technologies and interventions that enable pregnancies to be recognized earlier in gestation, and thereby enable behavioral change.

References

- Anna Aizer, Laura Stroud, and Stephen Buka. Maternal stress and child well-being: Evidence from siblings. manuscript, Brown University, February 2009.
- Douglas Almond and Bhashkar Mazumder. Health capital and the prenatal environment: The effect of ramadan observance during pregnancy. *American Economic Journal: Applied Economics*, 3(4), October 2011.
- Douglas Almond, Lena Edlund, and Mårten Palme. Chernobyl's subclinical legacy: Prenatal

¹⁸See our discussion of this paper in section 2.3

- exposure to radioactive fallout and school outcomes in Sweden. *Quarterly Journal of Economics*, 124(4):1729–1772, 2009.
- S.H. Alwasel, Z. Abotalib, J.S. Aljarallah, C. Osmond, S.M. Alkharaz, I.M. Alhazza, G. Badr, and D.J.P. Barker. Changes in placental size during ramadan. *Placenta*, 31(7):607 – 610, 2010.
- Fereidoun Azizi, Hossein Sadeghipour, Behnam Siahkollah, and Nasrollah Rezaei-Ghaleh. Intellectual development of children born to mothers who fasted during pregnancy. *International Journal of Vitamin and Nutrition Research*, 74(5):374–380, September 2004.
- Elizabeth Cascio and Douglas Staiger. Skill, standardized tests, and fadeout interventions. working paper, Dartmouth Economics, May 2011.
- Raj Chetty, John N. Friedman, Nathaniel Hilger, Emmanuel Saez, Diane Whitmore Schanzenbach, and Danny Yagan. How does your kindergarten classroom affect your earnings? evidence from project star. *Quarterly Journal of Economics*, 2011. forthcoming.
- Janet Currie. Healthy, wealthy, and wise: Is there a causal relationship between child health and human capital development? *Journal of Economic Literature*, XLVII(1):87–122, March 2009.
- Janet Currie and Rosemary Hyson. Is the impact of shocks cushioned by socioeconomic status? the case of low birth weight. *American Economic Review*, 89(2):245–250, May 1999.
- E. Dikensoy, O. Balat, B. Cebesoy, A. Ozkur, H. Cicek, and G. Can. The effect of Ramadan fasting on maternal serum lipids, cortisol levels and fetal development. *Archives of gynecology and obstetrics*, 279(2):119, 2009.
- Will Dobbie and Roland G. Jr Fryer. Are high quality schools enough to increase achievement among the poor? evidence from the harlem children’s zone. *American Economic Journal: Applied Economics*, 3(3):158–187, 2011.
- Gabriele Doblhammer and James W. Vaupel. Lifespan depends on month of birth. *Proceedings of the National Academy of Sciences*, 98(5):2934–2939, February 2001.
- Orla Doyle, Colm P. Harmon, James J. Heckman, and Richard E. Tremblay. Investing in Early Human Development: Timing and Economic Efficiency. *Economics and Human Biology*, 7(1):1–6, 2009.

- PM Eaton and PA Wharton. *Applied Nutrition*, chapter Sorrento Study of Nutrition in Pregnancy, pages 89–91. John Libbey, London, 1982.
- Shahul H. Ebrahim, R. Louise Floyd, Robert K. Merritt II, Pierre Decoufle, and Deborah Holtzman. Trends in Pregnancy-Related Smoking Rates in the United States, 1987-1996. *JAMA*, 283(3):361–366, 2000.
- Erica Field, Omar Robles, and Maximo Torero. Iodine deficiency and schooling attainment in tanzania. *American Economic Journal: Applied Economics*, 1(4):140–169, October 2009.
- Peter Gluckman and Mark Hanson. *The Fetal Matrix: Evolution, Development and Disease*. Cambridge University Press, Cambridge, 2005.
- James Heckman, Lena Malofeeva, Rodrigo Pinto, and Peter Savelyev. Understanding the mechanisms through which an influential early childhood program boosted adult outcomes. manuscript, University of Chicago, August 2010.
- James J. Heckman. The economics, technology, and neuroscience of human capability formation. *PNAS*, 104(33):13250–13255, August 14 2007.
- James J. Heckman and Dimitriy V. Masterov. The productivity argument for investing in young children. *Review of Agricultural Economics*, 29(3):446–493, 2007.
- ES Hunter and TW Sadler. D-(-)-beta hydroxybutyrate-induced effects on mouse embryos in vitro. *Teratology*, 36:259–64, 1987.
- Thomas Kane and Douglas Staiger. The promise and pitfalls of using imprecise school accountability measures. *The Journal of Economic Perspectives*, 16(4):91–114, Autumn 2002.
- Amita Kapoor, Elizabeth Dunn, Alice Kostaki, Marcus H. Andrews, and Stephen G. Matthews. Fetal programming of hypothalamo-pituitary-adrenal function: Prenatal stress and glucocorticoids. *Journal of Physiology*, 572(1):31–44, 2006.
- Elaine Kelly. The scourge of asian flu: in utero exposure to pandemic influenza and the development of a cohort of british children. *The Journal of Human Resources*, 46(4): 669–694, 2001.
- H. Kieler, O. Axelsson, S. Nilsson, and U. Waldenstroem. The length of human pregnancy as calculated by ultrasonographic measurement of the fetal biparietal diameter. *Ultrasound in Obstetrics and Gynecology*, 16(5):353–357, 1995.

- Alan B. Krueger. Experimental estimates of education production functions. *The Quarterly Journal of Economics*, 114(2):497–532, 1999.
- Alan B. Krueger and Diane M. Whitmore. The effect of attending a small class in the early grades on college-test taking and middle school test results: Evidence from project star. *The Economic Journal*, 111(468):1–28, 2001.
- Kaja Z LeWinn, Laura R Stroud, Beth E Molnar, James H Ware, Karestan C Koenen, and Stephen L Buka. Elevated maternal cortisol levels during pregnancy are associated with reduced childhood iq. *International Journal of Epidemiology*, 38(6):1700–1710, 2009.
- Ashok Malhotra, P.H. Scott, J.S. Scott, H. Gee, and B.A. Wharton. Metabolic changes in asian muslim pregnant mothers observing the ramadan fast in britian. *British Journal of Nutrition*, 61(3):663–672, May 1989.
- AM Prentice, A Prentice, WH Lamb, PG Lunn, and S. Austin. Metabolic consequences of fasting during ramadan in pregnant and lactating women. *Human nutrition. Clinical nutrition*, 37(4):283–94, July 1983.
- T Rizzo, BE Metzger, WJ Burns, and K Burns. Correlations between antepartum maternal metabolism and child intelligence. *New England Journal of Medicine*, 325:911–916, 1991.
- Jesse Rothstein. Teacher quality in educational production: Tracking, decay, and student achievement. *The Quarterly Journal of Economics*, 125(1):175–214, 2010.
- Anna Maria Siega-Riz, Tracy Hermann, David A Savitz, and John M Thorp. Frequency of eating during pregnancy and its effect on preterm delivery. *American Journal of Epidemiology*, 153(7):647–652, 2001.
- Reyn Van Ewijk. Long-term health effects on the next generation of ramadan fasting during pregnancy. *Journal of Health Economics*, 30:1045–1058, 2011.

Table 1: Summary Statistics

		“Muslims”	Caribbeans	White British	Indians	Other
Math	Mean	-0.36	-0.20	0.04	0.05	-0.08
	(SD)	(1.04)	(0.99)	(0.99)	(0.97)	(1.03)
	N	221,873	106,543	4,426,857	106,543	770,412
Reading	Mean	-0.38	-0.14	0.04	0.05	-0.10
	(SD)	(1.01)	(1.00)	(0.99)	(0.92)	(1.05)
	N	221,855	106,538	4,426,772	116,612	770,350
Writing	Mean	-0.31	-0.13	0.03	0.09	-0.09
	(SD)	(1.06)	(1.02)	(0.99)	(0.94)	(1.05)
	N	221,856	106,540	4,426,696	116,612	770,333
First Principal Component	Mean	-0.60	-0.28	0.06	0.11	-0.15
	(SD)	(1.66)	(1.59)	(1.55)	(1.49)	(1.65)
	N	221,818	106,522	4,426,410	116,602	770,132
Free School Meal	Mean	0.35	0.34	0.16	0.11	0.26
	(SD)	(0.48)	(0.47)	(0.36)	(0.32)	(0.44)
	N	221,902	106,555	4,426,973	116,625	770,569

Note: The first principal component is based on a principal components analysis of math, reading, and writing.

Table 2: Effects of Prenatal Ramadan Exposure on KS 1 Scores of Muslims

*Coefficients on Muslim * Ramadan Exposure*

Month Ramadan Began	<i>Dependent Variable</i>						
	Math (1)	Reading (2)	Writing (3)	PCA (4)	Math (5)	Math (6)	FSM (7)
<i>Probably</i>	-0.003	-0.051 **	-0.031	-0.051	-0.004	-0.019	0.027 **
<i>Not Exposed</i>	(0.026)	(0.024)	(0.026)	(0.040)	(0.022)	(0.012)	(0.012)
0	-0.054 **	-0.049 **	-0.051 **	-0.089 **	-0.039 **	-0.031 *	0.022 **
<i>(conceived)</i>	(0.023)	(0.022)	(0.023)	(0.036)	(0.02)	(0.016)	(0.011)
1	-0.068 **	-0.054 **	-0.052 *	-0.100 **	-0.045 *	-0.046 **	0.011
	(0.028)	(0.028)	(0.028)	(0.044)	(0.024)	(0.018)	(0.013)
2	-0.059 *	-0.067 **	-0.053 *	-0.103 **	-0.057 **	-0.048 **	0.006
	(0.032)	(0.032)	(0.032)	(0.050)	(0.028)	(0.019)	(0.015)
3	-0.081 **	-0.073 **	-0.055 *	-0.120 **	-0.063 **	-0.035 **	-0.004
	(0.033)	(0.032)	(0.032)	(0.051)	(0.028)	(0.017)	(0.015)
4	-0.046	-0.038	-0.036	-0.069	-0.047 *	-0.020	-0.002
	(0.029)	(0.029)	(0.030)	(0.047)	(0.026)	(0.015)	(0.014)
5	-0.023	-0.021	-0.023	-0.039	-0.032	-0.007	-0.003
	(0.028)	(0.027)	(0.028)	(0.044)	(0.024)	(0.014)	(0.013)
6	-0.022	-0.025	-0.011	-0.034	-0.016	-0.012	0.007
	(0.026)	(0.026)	(0.027)	(0.041)	(0.023)	(0.014)	(0.012)
7	-0.037	-0.033	-0.024	-0.054	-0.025	-0.003	0.026 **
	(0.025)	(0.025)	(0.025)	(0.040)	(0.022)	(0.012)	(0.012)
8	-0.027	-0.032	-0.016	-0.042	-0.025	-0.002	-0.001
	(0.023)	(0.023)	(0.023)	(0.036)	(0.02)	(0.011)	(0.011)
9	-0.024	-0.036 *	-0.040 *	-0.058 *	-0.015	0.005	-0.007
<i>(born)</i>	(0.021)	(0.020)	(0.021)	(0.032)	(0.018)	(0.015)	(0.010)
<i>Output Area FE's</i>	Yes	Yes	Yes	Yes	No	Yes	Yes
<i>Diff in Diff</i>	Yes	Yes	Yes	Yes	Yes	No	Yes
<i>N</i>	326,549	326,526	326,529	326,549	326,549	220,844	326,592

Note: Each column is a separate regression. Columns 1 to 5 and 7 include Caribbeans and show the interaction of the exposure measure with a dummy for Muslim. The coefficients on Caribbeans are shown in Table 3. Column 6 only uses Muslims. PCA refers to the first principal component of math, reading and writing. FSM refers to free school meal status.

* p<0.10, **p<0.05, *** p<0.01

Table 3: Prenatal Ramadan Exposure on KS 1 Scores of Caribbeans

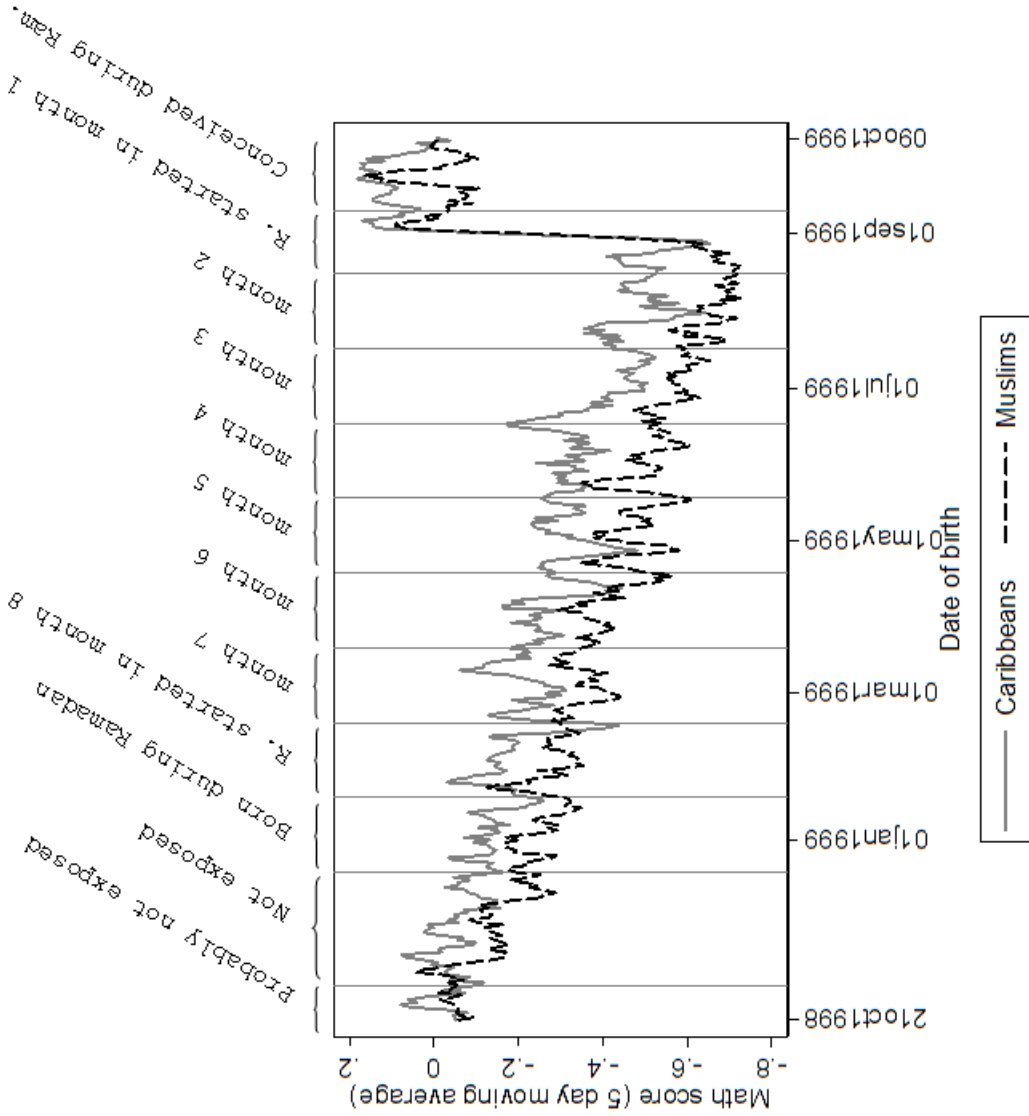
Coefficients on Ramadan Exposure for Caribbeans

Month Ramadan Began	<i>Dependent Variable</i>						
	Math (1)	Reading (2)	Writing (3)	PCA (4)	Math (5)	Math (6)	FSM (7)
<i>Probably</i>	0.007	0.042 **	0.016	0.040	0.018	–	-0.020 **
<i>Not Exposed</i>	(0.021)	(0.020)	(0.021)	(0.033)	(0.018)		(0.010)
0 <i>(conceived)</i>	0.038 *	0.043 **	0.036 *	0.067 **	0.025	–	-0.007
	(0.020)	(0.019)	(0.019)	(0.031)	(0.016)		(0.009)
1	0.040 *	0.036	0.038	0.065 *	0.018	–	-0.005
	(0.023)	(0.023)	(0.023)	(0.037)	(0.02)		(0.011)
2	0.012	0.027	0.008	0.027	0.004	–	0.006
	(0.026)	(0.027)	(0.026)	(0.042)	(0.023)		(0.013)
3	0.036	0.040	0.007	0.047	0.014	–	0.014
	(0.027)	(0.027)	(0.027)	(0.043)	(0.023)		(0.013)
4	0.010	0.010	0.004	0.013	0.010	–	0.011
	(0.024)	(0.025)	(0.025)	(0.039)	(0.021)		(0.012)
5	0.005	-0.004	-0.011	-0.007	0.013	–	0.010
	(0.023)	(0.023)	(0.023)	(0.037)	(0.020)		(0.011)
6	0.013	0.007	-0.012	0.005	0.007	–	-0.000
	(0.022)	(0.022)	(0.022)	(0.034)	(0.019)		(0.010)
7	0.028	0.019	0.004	0.028	0.009	–	-0.016
	(0.021)	(0.021)	(0.022)	(0.033)	(0.018)		(0.010)
8	0.027	0.020	0.002	0.028	0.021	–	-0.003
	(0.020)	(0.019)	(0.020)	(0.031)	(0.017)		(0.009)
9	0.023	0.026	0.023	0.042	0.011	–	-0.000
<i>(born)</i>	(0.017)	(0.017)	(0.017)	(0.027)	(0.015)		(0.008)
<i>Output Area FE's</i>	Yes	Yes	Yes	Yes	No	Yes	Yes
<i>Diff in diff</i>	Yes	Yes	Yes	Yes	Yes	No	Yes
<i>N</i>	326,549	326,526	326,529	326,549	326,549	220,844	326,592

Note: Each column is a separate regression and the columns correspond to those shown in Table 2. Entries show the coefficients on Ramadan exposure among Caribbeans. Column 6 has no entries because only Muslims are included in the regression. PCA refers to the first principal component of math, reading and writing. FSM refers to free school meal status.

* p<0.10, **p<0.05, *** p<0.01

Figure 1: Average Math Scores by Date of Birth and Timing of Ramadan Exposure for one Islamic Year



Note: Figure shows 5-day moving average Math scores for Muslims and Caribbeans born in one Islamic year (Oct. 21, 1998 to Oct. 9, 1999). The horizontal axis shows dates of birth. In the depicted year, Ramadan started on Dec. 19, and ended on Jan. 17. Vertical sections indicate periods of pregnancy during which exposure could have taken place. Months are calculated as 30-day periods. Note that the period “Ramadan started in month 1 of pregnancy” is a few days shorter than the periods “Ramadan started in month 2-9 of pregnancy”. The reason is that an average pregnancy takes 266 days, which is somewhat less than nine (30-day) months. Calculating backwards, starting from the date of birth, the “month 1”-group consists of 26 days.

Figure 2: Share of Pakistani/Bangladeshi Students by Local Authority, England 1998-2007

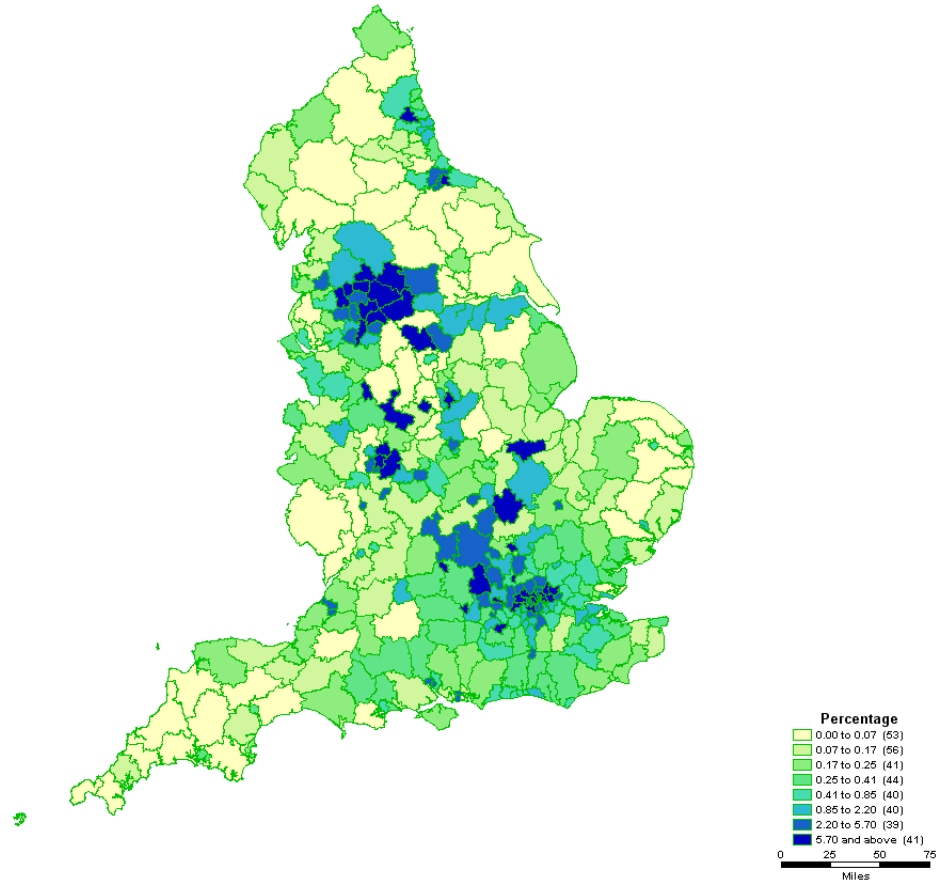
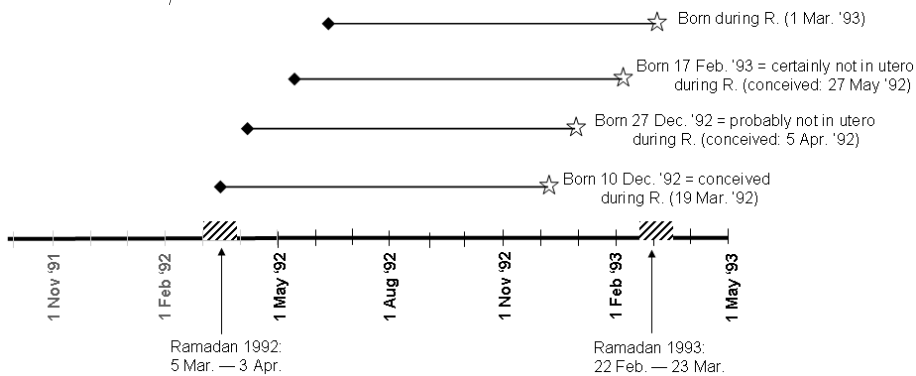
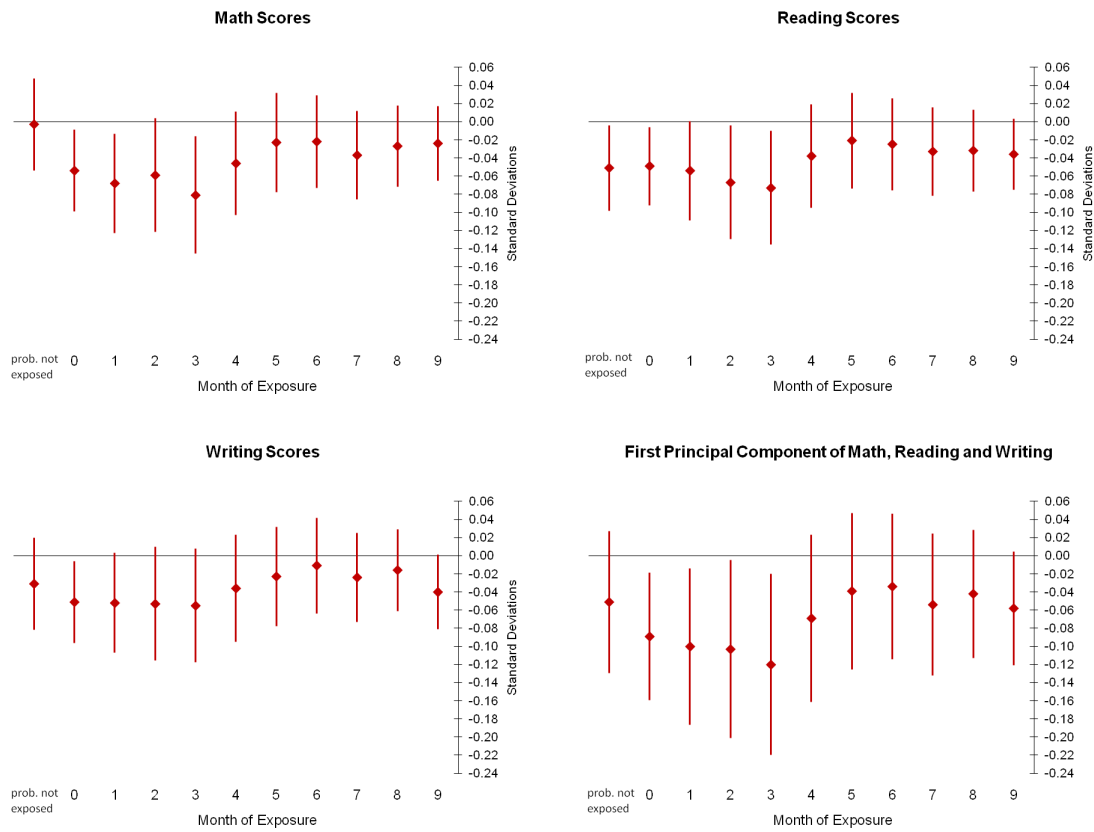


Figure 3: Calculating whether a person was in utero during a Ramadan Example: people born in 1992/3



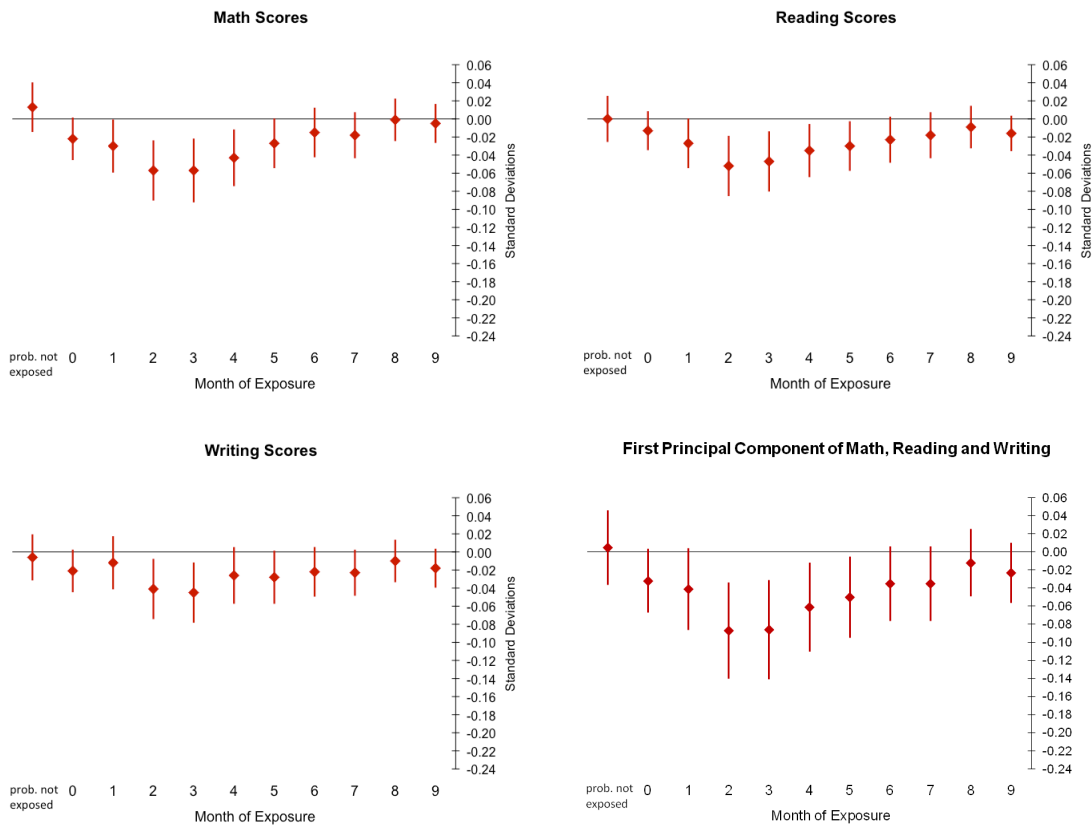
Note: Figure shows people born between 1 Nov. 1991 and 1 May 1993. Stars indicate birthdates, diamonds the calculated day of conception. Each line is 266 days long (the average length of human gestation). The shaded areas indicate Ramadans.

Figure 4: Prenatal Ramadan Exposure and Key Stage 1 Scores: Difference-in-Differences Muslims vs Caribbeans



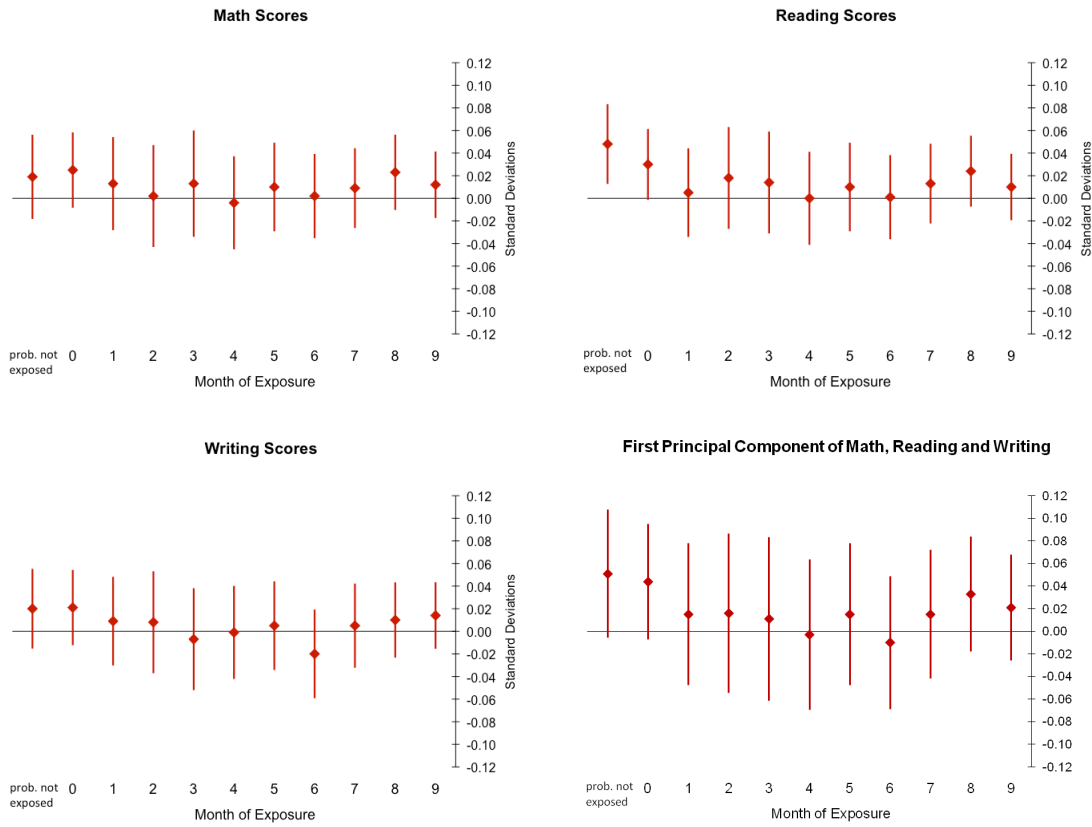
Note: Figures show 95 percent confidence intervals for the interactions of Ramadan exposure with a dummy for Muslim, from difference-in-differences estimates for Muslims vs Caribbeans (see Table 2). Math, Reading and Writing are in units of z-scores. The first principal component of math, reading and writing has a standard deviation of 1.58.

Figure 5: Prenatal Ramadan Exposure and Key Stage 1 Scores: Difference-in-Differences Muslims vs White British



Note: Figures show 95 percent confidence intervals for the interactions of Ramadan exposure with a dummy for Muslim, from difference-in-differences estimates for Muslims vs White British. Total sample size for each analysis is 4.6 million. Math, Reading and Writing are in units of z-scores. The first principal component of math, reading and writing has a standard deviation of 1.58.

Figure 6: Prenatal Ramadan Exposure and Key Stage 1 Scores: Difference-in-differences
Caribbeans vs White British



Note: Figures show 95 percent confidence intervals for the interactions of Ramadan exposure with a dummy for Caribbean, from difference-in-differences estimates for Caribbeans vs white British which is used as a placebo test. Math, Reading and Writing are in units of z-scores. The first principal component of math, reading and writing has a standard deviation of 1.58.