HEDG
HEALTH, ECONOMETRICS AND DATA GROUP

THE UNIVERSITY of York

WP 12/02

Intergenerational effect of schooling and childhood overweight

Ryota Nakamura

January 2012

york.ac.uk/res/herc/hedgwp

Intergenerational effect of schooling and childhood overweight

Ryota Nakamura[*][†]

January 18, 2012

Abstract

Prevalence of overweight among children is at the top of health policy agenda in many developed countries. We study the causal effect of mothers' schooling on children's body weight. We exploit the 1972 schooling reform in England and Wales, which raised the minimum school leaving age from fifteen to sixteen. Our regression-discontinuity estimates use Health Survey for England (1998-2002) and show that the extra year of schooling for mothers induced by the reform significantly reduces their son's weight. There is only insignificant negative effect for daughters. Additionally, we do not find that mothers' schooling improves children's health behaviour (fruit and vegetable consumption; exercising).

Keywords: Overweight children, Schooling, Regression-discontinuity.

JEL: I12, I20.

---

[*]School of Economics; Norwich Medical School, University of East Anglia, Norwich, NR4 7TJ, United Kingdom. E-mail: R.Nakamura@uea.ac.uk.

[†]The Health Survey for England is provided by UK Data Archives. I thank Luigi Siciliani and Andrew Jones for their helpful comments and suggestions. All errors are my own.

1 Introduction

Childhood overweight has been one of the most serious public health problems in developed and developing societies. In England, the prevalence of overweight children has increased at an alarming rate since the late 1980s. Rona and Chinn (2001) report that it was 5.4 percent for boys and 9.3 percent for girls in 1984, rising to 9.0 percent for boys and 13.5 percent for girls in 1994. Lobstein et al. (2004) update the information and note that the rate reached as high as 17.0 percent for boys and 23.6 percent for girls in 1998.[1]

The cost of overweight among children can be considerable. It includes health care cost in relation to associated diseases such as hypertension, cardiovascular disease and depression. Studies have found that overweight in childhood is also an important risk factor of such diseases in their adulthood (for example, Ebbeling et al. 2002). Also, overweightness may lead to lower educational achievements, and hence lower economic outcomes in the future (Ding et al., 2006). Preventing overweight among children may also reduce overweight (or obesity) among adults because overweight children are more likely to become overweight adults.

In this study we investigate the relationship between a mother's schooling and her children's weight measures. In particular, we examine if the relationship is causal. Mothers' education is one of the most important socioeconomic factors predicting children's health (Case et al., 2002; Currie, 2009). Previous studies confirm that the children of educated mothers are more likely to be healthier as educated mothers tend to have higher salaries or marry high-earning husbands; they can therefore afford healthy foods. They are also better at processing nutritional information about various foods and are therefore more efficient at managing their children's weight.

Little is known, however, about whether mothers' schooling has a causal effect on children's weight. It is unlikely that children's weight affects mothers' schooling.

[1] Rona and Chinn (2001) use a sample of children aged 4-11, while Lobstein et al. (2004) use children aged 7-11. They both use the definition of overweight which is recommended by the International Obesity Task Force.

However, unobserved third factors, such as a mother's general ability, will power, and time preference, may explain both her own schooling and children's weight. We investigate the causality by exploiting the 1972 reform of the compulsory schooling law in England and Wales, which provides an exogenous variation in mothers' schooling.[2] In the English and Welsh education system, all children of compulsory schooling age must engage in full time education. The compulsory schooling age is regulated by the law. In 1972, the minimum school leaving age was raised from fifteen to sixteen by the Raising of School Leaving Age (RoSLA) policy. That is, children who were born before September 1957 were permitted to leave school and work at age fifteen, whereas those who were born after this date were forced to stay in school until they became sixteen.

We employ regression-discontinuity design (RDD) to examine if the additional schooling induced by the reform has had a significant impact on the next generation's bodyweight. RDD is close to a local randomised experiment (Lee and Lemieux, 2010). We regard mothers who were born just before September 1957 (control group) as the counterfactual of those who were born just after this cut-off point (treatment group). The effect of the reform should be viewed as a local average treatment effect to mothers who would have left school at age fifteen, rather than the average treatment effect to the general population. We therefore focus on mothers who left school at age fifteen or sixteen in the main analysis (Black et al., 2005; Lindeboom et al., 2009). Finally, previous studies which exploit schooling reforms mostly use year of birth as the assignment variable to determine the treatment status (Doyle et al., 2007; Lindeboom et al., 2009). However, since schooling reforms (including the 1972 reform) were often implemented in the middle of the year, measurement error in the treatment status may bias their results (Lee and Card, 2008). We use month of birth as well as year of birth as the assignment variable. We thus identify the treatment status much more efficiently.

[2] We focus to provide some clucial aspects of the reform only in the paper. More detailed review of the reform is available in the literature (for example, see Del Bono and Galindo-Rueda, 2007; Anderberg and Zhu, 2010; Clark and Royer, 2010).

Our main result is based on children aged five to fifteen and their mothers who were born between September 1952 and August 1962 and left school at age fifteen or sixteen from Health Survey for England (1998-2002). We find that the extra year of schooling generated by the reform significantly reduces their sons' weight. There is little effect for daughters. The result is overall robust to different model specifications and estimation strategies. In addition to the main analysis, we investigate the impact of mothers' schooling on children's health behaviour such as fruit and vegetable consumption and exercising.

The paper proceeds as follows. Section 2 reviews previous literature. Section 3 describes our data set. Section 4 explains the empirical framework. Section 5 shows the main results. Section 6 gives the results of additional analyses. Finally, Section 7 concludes.

2 Literature

The association between parent's schooling and children's health is well documented in the literature (Currie, 2009 for survey). A large and significant association between mother's educational attainment and children's general health status is found in the US (Case et al., 2002), Canada (Currie and Stabile, 2003), and England (Currie et al., 2007). Chen and Li (2009) use samples from China, and find that mothers' education is positively associated with adopted children's height-for-age Z-score.[3] They conclude that the positive association between mothers' education and children's health is not merely due to genetic factors. Finally, Stifel and Averett (2009) use samples from NLSY79 in the US, and find that mothers' education is associated with a decline of overweight for white boys, and also a decline of underweight for white girls.

Several studies address the endogeneity in schooling (Currie and Moretti, 2003; McCray and Royer, 2006; Chou et al., 2010; Doyle et al., 2007; Lindeboom et al.,

[3] The height-for-age Z-score is defined as the standard deviation point of one's height compared to the reference distribution from the general population. Details of derivation is given in the next section.

2009).[4] Among these studies, the following two studies are closest to our own study. First, Doyle et al. (2007) investigate the effect of parents' income and education on children's self-assessed health and chronic conditions. They use samples from the Health Survey for England (1997-2002). Their instrumental variable approach uses the 1972 compulsory education reform, which raised the minimum school leaving age from fifteen to sixteen. They allow for the impact of the reform on schooling varying across regions and find that it is significant in the North West region, where historically people are less educated. They find that mothers' schooling reduces children's self-assessed health and chronic conditions. Compared to our study, they use the same data set (HSE) and the reform (the 1972 reform). However, our study is different from theirs in several ways. First, we examine childhood overweight. We investigate variables such as body mass index, fruit and vegetable consumption, and physical activity. Second, we use regression-discontinuity framework to compare between parents who were born just before and just after the cut-off point. Finally, Doyle et al. (2007) highlight the regional variation in the impact of the reform on at all levels of education. By contrast, we focus on those who left school at age fifteen or sixteen (thus those who were directly affected by the reform) to verify the significance of the reform.

Lindeboom et al. (2009) examine the relationship between parents' schooling and children's various health outcomes including BMI and overweight status. Their sample is derived from children who were born in 1958 and part of the National Child Development Study in the UK. They use the 1947 compulsory education reform, which raised the minimum school leaving age from fourteen to fifteen, to instrument parents' schooling. They find that parents' schooling has little effect on all health outcomes of the children they study. However, they find that schooling significantly reduces financial difficulties. Their study is close to our study in that it investigates children's weight for the same country (UK). However, their analysis is based on

[4] There are larger number of studies which investigate the causal relationship between education and adults' own health or own weight (Lleras-Muney, 2005; van Kippersluis et al., 2009; Clark and Royer, 2010; Blunello et al., 2009; Webbink et al., 2010). Their conclusions are mixed.

children who were born in 1958. As childhood overweight has been prevalent since the late 1980s, it is important to extend their study by using an updated data set. We use a different reform which was conducted in 1972. We analyse children living between 1998 and 2002, which is reasonably recent. As shown later, we present the opposite result to their analysis. We find that the additional schooling has a significant effect on children's weight. Moreover, we examine children's health behaviour (fruit and vegetable consumption and exercising).

3 Data

We use the Health Survey for England 1998-2002. The HSE is an annually repeated cross-sectional survey, which has been conducted since 1992. Using the Postcode Address File, national representative samples from private households are studied.[5] Originally, the HSE covered adults only. From the 1995 survey onwards, up to two children from each household are studied.[6] The survey collects data by face-to-face interviews, self-completed questionnaires and nurse visits. The survey covers various topics including demographic, socioeconomic and health status.

In the HSE, information about month of birth and year of birth is available. Month of birth is available in 1998-2002 surveys only. Identifying the month of birth is crucial because it precisely determines whether a mother was exposed to the reform or not.

The sampling framework of the HSE is at the household level, therefore we can match the information on a child with his/her mother's characteristics.[7] The estimation sample is selected in the following way. We focus on school children aged five to fifteen. We use mothers who were born between September 1952 and August

[5] In each year the HSE oversamples particular groups of society (such as elders, ethnic minorities) in addition to general population. This study uses a general population sample only.

[6] Two children are randomly selected. The HSE defines children as individuals who are fifteen years old or younger. With children aged less than twelve, parents answer the questions with the child present. Children aged thirteen to fifteen can answer the questionnaire by themselves. Finally, adults (older than sixteen) answer by themselves.

[7] We do not distinguish between natural, adopted, foster and step mothers.

1962 to focus on those who were born just before or just after the reform. We follow previous studies to choose this 10-year sample window. We later check the robustness of the result by re-estimations using different sample windows (the results are shown in Table 3 and Table 5). We simply drop observations which contain item non-response.

We measure children's weight by BMI for age and gender Z-score (hereafter BMI Z-score). BMI is derived using measured height and weight. The Z-score is calculated by the LMS method using the reference distribution for British children in 1990.[8] Descriptive statistics are shown in Table 1.

[**Table 1 here**]

We show the results separately by mother's treatment status (i.e. pre-reform and post-reform cohorts) and by gender of the children.[9] We provide results from two different samples. The first is the full sample, where mothers are from all education levels. The second is the restricted sample, where we use mothers who left school at age fifteen or sixteen. This is because the reform only affected those who would have left school at age fifteen. For both full and restricted samples, we use mothers who were born between September 1952 and August 1962 (10 years window). 'Pre-reform' mothers were born before September 1957, whereas 'post-reform' mothers

[8] The LMS method presents children's BMI distribution at each age by its median (M), standard deviation (S), and skewness based on the Box-Cox power (L). L, M and S differ by age and gender. Using the information, it transforms the raw BMI distribution to a normal distribution. The BMI Z-score is defined as the standard deviation point of the reference normal distribution. It is computed by the following equation:

$$\text{BMI Z-score} = \frac{[\text{BMI}/M(t)]^{L(t)} - 1}{L(t)S(t)}.$$

Here t represents the age in years. For more detail, see Cole et al. (1995).

[9] 'Pre-reform' mothers are older than 'post-reform' mothers, and the samples of their children (aged five to fifteen) are all taken in 1998-2002 surveys. Therefore, we may oversample the children of 'post-reform' mothers because children of 'pre-reform' mothers are more likely to be older than fifteen in 1998-2002. Although including mother's age variable into the model does not seriously affect the results, and also our results are not sensitive to the choice of children's age band, we are not able to avoid this possible sample selection problem. Moreover, Black et al. (2008) find that the schooling reform reduces teenage pregnancy.

were born after this date. The overall results are robust to different sample windows (this is formally examined in the result section). Column A shows the BMI Z-score and overweight rate of children in 1998-2002 from the full sample.[10] The mean BMI Z-score is around 0.5-0.6 for both genders. This implies that children in 1998-2002 are heavier by 0.5-0.6 standard-deviation points than they were in 1990. Overweight rate ranges between 23 and 30 percent. Compared to the post-reform sample, the pre-reform sample exhibits a slightly higher BMI Z-score and overweight rate for sons, but it shows a lower BMI Z-score and overweight rate for daughters. This odd result for daughters may imply that the reform decreased son's weight while there was no effect for daughters. Next, column B shows the result from the restricted sample. Both sons and daughters are slightly heavier compared to the full sample case. This may because children of less educated mothers are more likely to be heavier. Compared to the post-reform sample, the pre-reform sample displays a higher BMI Z-score and overweight rate for sons, while it displays a lower BMI Z-score and overweight rate for daughters.

We measure mothers' schooling by the age they left school.[11] In the HSE, the variable is continuous but is bottom-coded at fourteen and top-coded at nineteen (i.e. 14 or less, 15, 16, 17, 18, 19 or more). Bottom-coding may not be so serious because the number of mothers who left school at fourteen or below is very small (1.7 percent). However, top-coding can be serious. As for mothers who left school at nineteen or above, we recode the variable following Doyle et al. (2007).[12] The

[10] Overweight rate is derived from BMI Z-scores. The definition of overweight for children varies by age and gender because their BMI changes substantially with age and gender (BMI increases more steeply for infants and moderately for young adult; and the rate also differs across genders). In this study we use the threshold value of BMI Z-scores, which is given by Cole et al., (2000), to compute the overweight rate.

[11] Mothers are asked to answer the following question. "At what age did you finish your continuous full-time education at school or college?"

[12] We proceed as follows. Health Survey for England contains the information about one's highest academic degrees. We combine this information with UK Labour Force Survey (LFS) to recode the average school laving age according to one's highest qualification. In the LFS survey, the average school leaving age of those who have a university qualification is twenty-one. Therefore school

average school leaving age is 17.12 (See Table A1 in the Appendix).

In addition to these key variables, we have other control variables. They are mother's age, children's age, region of residence dummies, and survey year dummies. The average age of mothers is 43.9. The average age of sons is 10.74, whereas the average age of daughters is 10.70. The detail is given in Table A1 in the Appendix.

4 Econometric framework

The purpose of this study is to establish causal relationship between mothers' schooling and children's weight. We employ a regression-discontinuity design (Hahn et al., 2001; Lee and Lemieux, 2010). The reason for this is as follows. The regression-discontinuity estimates have better internal validity compared to other estimation strategies such as instrumental variable approach. This is mainly due to the fact that the regression-discontinuity design focuses on the identification of causal effects around the cut-off point (the 1972 reform). The RDD estimates thus provides with the weighted Wald estimates, in which the weight is the probability of being born around the cut-off point (Imbens and Lemieux, 2008).[13]

We estimate the causal effect of mothers' schooling on children's BMI Z-scores. The 1972 reform provides a sudden increase in mothers' schooling, particularly for

leaving age is recoded to twenty-one if one has a university degree. Also, the average leaving age for those who have teaching qualification is twenty. Therefore we recode the school leaving age to twenty if one has a teaching qualification. As a result, school leaving age ranges between fourteen and twenty-one.

[13] The crucial assumptions for the identification in regression discontinuity design are as follows. First, the treatment status (i.e. whether one is exposed to the reform or not) is not randomly assigned, but deterministically assigned. In this study, September of 1957 is the deterministic cut-off point. Second, individuals cannot manipulate the assignment variable (i.e. the birth month) precisely. Third, other variables in the econometric specification (which are described below: children's age, mothers' age, place of birth) should be smooth functions of the assignment variable at the cut-off point. This is tested by a simple approach suggested by Lee and Lemieux (2010). We test if there is any discontinuity in each of control variables at the cut-off by regressing these variables on the reform dummy. As a result, no discontinuity is found for all variables.

those who would have left school at age fifteen.[14,15] Month of birth and year of birth determine whether a mother is exposed to the reform or not. Specifically, those who were born after September 1957 were exposed to the reform. Compliance to the reform was not universal and many mothers left school at age fifteen or below, even though they were exposed to the reform (about 9 percent of the full sample).[16] We therefore conduct a 'fuzzy' regression-discontinuity analysis.

As explained in the previous section, we focus on mothers who were born between September 1952 and August 1962 (10-year window) in order to avoid using samples which are far from the cut-off point. Also, we focus on children aged five to fifteen. The effects of mothers' schooling may be different for sons and daughters. All analyses are done separately by gender of children.

Let S be the mother's age when she left schooling, and W be children's weight.

[14] Note that in this study we focus on the jump in the schooling at September 1957 cohort, which was due to the nation-wide reform of minimum school leaving age. Some readers may wonder if the comprehensive education reform, which was implemented around this period, also affects the estimate. The comprehensive education reform have changed the 'quality' of education (not the length of the schooling), and it was conducted slowly at different places and different timings during 1960s and early 1970s. Therefore this reform is not likely to affect the discontinuity in schooling at the September 1957 cohort.

[15] Students were not necessarily allowed to leave school on their exact day of birth (either fifteenth or sixteenth). The timing that students were allowed to leave school depended on their birthday, and was set as follows. Those whose sixteenth birthday falls between 1st September and 31st January are allowed to leave school at the end of the spring term. Those whose sixteenth birthday falls between 1st February and 31st August are allowed to leave on the last Monday of May (Anderberg and Zhu, 2010). This means that, for students who wanted to leave school at their earliest opportunity, they were equally compelled to stay in school for additional one year by the reform irrespective of their birthday, but the timing to leave school during the year was different depending on their birthday.

[16] The main reason for this is that some of those who were born in late June, July and August left school at age of fifteen even after the reform was implemented (Clark and Royer, 2010). This may be because the academic year essentially ends before the end of August (exam periods finishes in mid-June). Therefore students who were born between late June and August could leave school before the minimum school leaving age of sixteen. In our data, only small number of mothers report that they left school at age of fourteen or less. This could be either due to misreporting or due to violation of the law. We are not able to judge. We decided to remove such samples from the estimation.

Also, let X represents the mother's birth cohort. X is a continuous variable, which is centred at September 1957 cohort (i.e. $X = 0$), who were exposed to the reform first. For estimation we follow the way which is suggested by Lee and Lemieux (2010). We estimate the following equations:

$$S_i = \alpha + \tau D_i + f(X_i) + Z_i'\theta + \varepsilon_i \quad (1)$$

$$W_i = \beta + \delta D_i + g(X_i) + Z_i'\phi + \epsilon_i. \quad (2)$$

First, we estimate Eq.(1). Now D indicates whether a mother was exposed to the reform or not. We control for birth cohorts to isolate the jump in schooling at the cut-off point from the long-term trend. The function $f(X)$ is the low-order polynomial function of X, which is fully interacted with the reform dummy (D) to allow for the functional form being different on either side of the cut-off point.[17] The additional control Z includes: mother's age, mother's age squared, children's age, children's age squared, region of residence dummies, and survey year dummies.[18,19] We expect these variables to reduce sampling variations. We check if the result is robust to the inclusion of these variables.[20] We show results with and without these additional variables.

[17] We estimate the treatment effect at the cut-off point, where X takes zero in Eq.(1) and Eq.(2). At the cut-off point, the interaction term disappears (or at least converges to zero) and only the main effect of D remains.

[18] We include children's age effect as an additional control variable. The age effect may capture the social influence on children's weight (such as school curriculum and exposure of fast-food restaurants; Powell and Bao, 2009). Note that we take into account the biological factors in the construction of BMI Z-score. As shown in the next section, including the age effect does not affect the main result substantially.

[19] Some other potential control variables may be considered, such as father's education, household income, and other family situation. However, these variables can be the outcome of mother's education, and therefore they would be endogenous. In this study we decide not to include such variables.

[20] In doing this, we follow Lee and Lemieux (2010). We estimate the first stage equation (without Z) by replacing the dependent variable with mother's age or child's age. We check if the reform affected these additional control variables. In fact, the cohort effect has only a weak correlation with those variables.

The polynomial function should fit the data well. We check the goodness-of-fit of the polynomial function using the G-test (Lee and Card, 2008).[21] Also, as X is a discrete variable, we estimate robust standard errors clustered at cohort level to take into account possible non-random specification errors at the cut-off point. We determine the order of the polynomial using Akaike Information Criterion (AIC). We proceed to the next stage if the model passes the G-test and it is suggested by AIC.

We move to examine the reduced-form effect of the 1972 reform on children's BMI Z-scores.[22] We do this by estimating Eq.(2) by OLS. We use the same specification as in the first stage estimation except for the dependent variable (Lee and Lemieux, 2010). We conduct a G-test to check if the polynomial function is sufficiently flexible. We estimate robust standard errors clustered at the cohort level. All estimations are done separately by gender of children.

Finally, we show the fuzzy RD estimate. This is interpreted as the effect of the extra year of schooling *induced by the 1972 reform* on children's BMI Z-scores. The effect is given by the ratio of the discontinuity in BMI Z-scores at the cut-off point to the discontinuity in mothers' schooling at the cut-off point (i.e. the Wald estimator). This is represented by δ/τ in Eq.(1) and Eq.(2). We estimate cluster robust standard errors at the cohort level.

We conduct a robustness check of the results using an alternative approach within RDD. We capture the discontinuity by employing a nonparametric function, rather than parametric polynomial function. Hahn et al (2001) and Porter (2003) suggest that local liner regression performs well in RDD settings. The treatment effect is

[21] The G-test compares between the explained sum of squares of the model with polynomial function (restricted model) and that of the model with a full set of dummy variables of X (unrestricted model). If the two are very different we reject the null hypothesis that the polynomial specification is sufficiently flexible.

[22] We conduct the same analysis using the probability of being overweight using the same sample. The results are overall consistent, and are available on request.

estimated by minimising the following objective function[23]:

$$\min \sum_i \mathbf{1}_{\{-h \leq X_i \leq h\}} \cdot (W_i - \alpha - \beta X_i - \tau S_i - \delta X_i \cdot S_i)^2.$$

Here h represents the bandwidth (or the sample window around the cut-off) that is used in the estimation. We use rectangular kernel for this estimation (Lee and Lemieux, 2010).

Bandwidth selection is crucial in this approach. Using smaller bandwidth enables us to estimate the precise effect at the cut-off, but this may risk of bias due to small number of observation. We obtain the optimal bandwidth by employing a 'leave one out' cross-validation procedure (Ludwig and Miller, 2007; Imbens and Lemieux, 2008). This procedure is as follows. We check how well the linear regression fits the data when we use the bandwidth h. We perform a linear regression without using a month-of-birth cohort (which the mother i belongs) and predict the value of $\hat{S}_i$ at $X_i = x_i$. We iterate this procedure for each single i in the bandwidth h. The optimal bandwidth h^* is derived by choosing h which minimises the mean square of the prediction errors.[24] We iterate the same procedure for the outcome variable (children's weight). In the actual analysis, we apply the same bandwidth for the schooling equation and the outcome equation (Imbens and Lemieux, 2008). If each procedure suggests different bandwidth, we use the shorter one in the estimation.

5 Results

In this section we discuss our main results, beginning with the first stage effect of the 1972 reform on mothers' schooling. Next, we show the reduced form estimate of the effect of the reform on children's BMI Z-scores. We then show the fuzzy regression discontinuity estimates of the effect of the extra year of schooling induced

[23] We follow the procedure suggested by Imbens and Lemieux (2008)

[24] More formally, the optimal bandwidth is such that:

$$h^* = \arg\min_h \frac{1}{N} \sum_i \left[S_i - \hat{S}_i(X_i) \right]^2.$$

by the reform on children's BMI Z-scores (with robustness check using local linear regression approach).

5.1 Effect of the 1972 reform on mother's schooling

In the first stage, we estimate the effect of the 1972 reform on mothers' schooling. First, we check if there is really a discontinuity in schooling at the cut-off point by using graphs. The 1972 reform affected only those who would have left school at fifteen. To take this into account, we use two sets of samples. The first is the full sample, where we use mothers from all education levels. The second is a restricted sample, where we use only mothers who left school at age fifteen or sixteen. Figure 1 and Figure 2 plot mothers' average school leaving age by monthly birth cohort.

[**Figure 1 and Figure 2 here**]

In Figure 1, we use the full sample. In Figure 2, we use the restricted sample. Figure 1 displays little discontinuity around the cut-off point. In unreported analysis, we find no statistically significant increase in schooling. The full sample is not suited for RD design and therefore we do not pursue it. By contrast, Figure 2 shows a discontinuity at the cut-off point. We superimpose the flexible quadratic fit onto the plots. The fitted lines suggest that the reform extended schooling by about 0.4 years. In the following we focus on the restricted sample (the restricted sample consists of mothers who left school at age of fifteen or sixteen only). Note that we no longer have the measurement error problem in schooling discussed in Section 3.

Given the graphical examination above, we formally analyse the effect of the 1972 reform on mothers' schooling. In the polynomial control approach, it is important that the polynomial function fits the data well. We check the goodness-of-fit of polynomial functions by G-test. The test indicates that the specification using single monthly cohort, as drawn in Figure 2, is too noisy. We decided to group ten monthly cohort to constract the assignment variable.[25] Next, we select the

[25] The procedure is as follows. To obtain a better fit of the function, we group several cohorts in order to increase the number of observations in each bin. We gradually increase the number

optimal order of polynomials by Akaike Information Criterion. We compute AIC of fully interacted linear to quartic specifications (not reported), and AIC suggests that we should adopt linear specification. In the following we present results from linear specification and also quadratic specification for robustness check. Finally, we check if the result is sensitive to the inclusion of additional controls. The additional control variables are mother's age, mother's age squared, children's age, children's age squared, region of residence dummies (nine regions in England), and survey year dummies. We can safely include these variables into the model as we do not find any discontinuity around the cut-off point in any of these variables.[26]

Table 2 displays the result of the effect of the reform on mothers' school leaving age.

[**Table 2 here**]

The reform raises mothers' schooling by about 0.35 to 0.45 years, and the effect is highly significant in all models.[27] The result is overall robust to the inclusion of additional controls. The size of the effect is not a full year. This is because 'pre-reform' mothers left school at age of fifteen or sixteen, whereas 'post-reform' mothers left school at age sixteen only. Finally, as explained above, the G-test indicates that the models are sufficiently flexible.

---

of observations in each bin until we pass the goodness-of-fit test. We iterate this process for four polynomial functions (from liner to quartic). The G-test is conducted for each polynomial order and the size of the bin. Finally, we pass the test when we include ten month-of-birth cohorts in each bin (non of the polynomial specifications from linear to quartic passed the test until then).

[26] As explained in the previous section, we estimate the first-stage equation by replacing the dependent variable with each additional control variable (Lee and Lemieux, 2010). By doing this we can simply examine if there is a discontinuity in each control variable around the cut-off point. We do not find any discontinuity in any of the variables (details not reported). Note that we do not mean to expect that the reform affect mother's birth age. This is to double-check the validity of the regression discontinuity design. Basically, we check if there is any jump in these control variables around the cut-off point. If there is any, it implies that the jump in the school leave age at September 1957 may not be due to the schooling reform, but may be due to other event.

[27] We check if the reform is a "weak" instrument. The F-statistic (not shown) is larger than 10 in all models, which implies that the reform is not weak (Staiger and Stock, 1997; Cameron and Trivedi, 2005).

We check the robustness of the effect by using narrower sample windows. In the main model, we use mothers who were born between September 1952 to August 1962, i.e. 60 month-of-birth cohorts on either side of the cut-off point. We now use 50 month-of-birth cohorts and 40 month-of-birth cohorts. As a result, the estimated effects are robust and are again highly significant. The magnitude of the coefficient is overall robust, too.

To summarise, we conclude that the 1972 reform has a significant impact on mothers' schooling. Note that the result is based on the restricted sample: all mothers left school at age fifteen or sixteen.

5.2 Effect of the reform on children's BMI Z-scores

Figure 3 and Figure 4 plot children's BMI Z-scores by mothers' month-of-birth cohort with flexible quadratic fit (separately by gender).

[**Figure 3 and Figure 4 here**]

Plots do not show a visually obvious discontinuity around the cut-off point in all cases. However, quadratic fit indicates that there is a small discontinuity at the cut-off point especially for sons' BMI Z-score.

Next, we formally examine the reduced form effect of the reform on children's BMI Z-scores. The results are given in Table 3.

[**Table 3 here**]

We use the same specification as in the first stage estimation. We conduct the goodness-of-fit test using the G-statistic. We find that all models are sufficiently flexible except for the linear polynomial cases for daughters. We are less confident in the estimates from these specifications; however, note that they still provide consistent estimates (Lee and Card, 2008).

Table 3 gives the result for children's BMI Z-scores. For sons, the results are sensitive to specification but are robust to the inclusion of additional controls. The effect is negative and statistically significant in all models. The size of the effect is relatively small (-0.2 and -0.25 points) if we control for linear polynomial function. The effect becomes larger (-0.41 and -0.45 point) if we control for quadratic function. The AIC and G-test show that quadratic specification fits better, hence it is preferred. For daughters, the effect can be positive and negative. The size of the effect is small (-0.24 to 0.09 point) and statistically insignificant.

5.3 Effect of the extra year of schooling on children's BMI Z-scores

We move to the result of fuzzy RD estimation. RD estimates represent the effect of the extra year of schooling induced by the reform on children's BMI Z-scores.[28] In addition to the results from the polynomial control approach we have used, we also present results from the local linear regression approach for robustness checks. Table 4 gives the result.

[**Table 4 here**]

For sons, the estimate implies that the extra year of schooling for mothers significantly reduces their sons' BMI Z-scores by about 0.55 to 0.9 points. The size of the effect is sensitive to specification, but it is overall large. When we use local linear regression approach, the estimated effect becomes larger and it amounts to 1 to 1.2 points. For daughters, the estimated coefficients are all statistically insignificant irrespective of model specification and econometric approach. However, the sign of the coefficient is negative (except for one case) as predicted.

To summarise, we find that a mother's schooling has a causal effect on her son's BMI Z-score. The magnitude of the effects is considerable. We do not find such effects for daughters. A possible reason why we find this gender discrepancy is that daughters may be more likely to be affected by their friends and 'fashion leaders.' This may be more pronounced for older girls as they are more likely to be conscious others. We are not able to examine this formally because of the limitation due to sample size.[29] Note that RDD just identifies the effect around the cut-off point.

[28] The implication is essentially the same if we analyse the probability of being overweight. The results is available on request.

[29] Interestingly, Black et al. (2005) also find a significant effect of parents' education for sons but

Second, we use the restricted sample: all mothers left school at age fifteen or sixteen. Third, RDD gives a local average treatment effect, where the estimate only reflects the effect of one additional year of schooling to mothers who complied with the reform.

6 Effect of the 1972 reform on children's health behaviour

In the previous section we found that additional schooling for mothers induced by the reform decreases the weight of their sons. In this section we investigate two possible pathways for this relationship: we investigate the effect of mothers' schooling on children's health behaviour. Specifically, we analyse (1) the total portion of fruit and vegetable consumption; and (2) whether children play any sports.[30] These factors are important because weight is determined by calorie intake and expenditure.

In the HSE, all respondents are asked how many portions of fruit and vegetable they had the day before the survey. Also, children are asked whether they have played any sports in the previous week (the variable is given as an indicator). Unfortunately, these variables are available only in limited survey years (2001 and 2002 for fruit & vegetable consumption; 1998, 1999, and 2002 for playing any sports), therefore the estimation sample is much smaller compared to that in the previous section. The analysis is based on the same specification as before. We focus on mothers who were born from September 1952 to August 1962. We restrict the sample to mothers who left school at age fifteen or sixteen. We analyse sons and daughters separately. The models pass relevant specification tests. To save space, we only present results without additional controls. The result is overall robust to

not for daughters.

[30] For fruit and vegetable consumption, children are asked to answer several questions such as "How many tablespoons of vegetables did you eat yesterday?" The HSE derives the variable "Total portion of fruit and vegetables," which we use in this study. For sports, children are asked to answer the question "In the last week, that is last (date last week) up to yesterday, have you/has (name of child) done any sports or exercise activities, not counting things done as part of school lessons?"

the inclusion of additional controls (detail not reported). Finally, to save space we focus on presenting RD effects (with polynomial control) only.

Block A of Table 5 gives the RD effect of mothers' schooling induced by the reform on children's fruit and vegetable consumption.

[**Table 5 here**]

For sons, the effect is positive (0.75 and 3.00) but statistically insignificant. For daughters, the effect is again positive but small and statistically insignificant. Block B displays the result for the probability of playing any sports in the previous week. For both sons and daughters, the effects of schooling are positive but statistically insignificant. This is somewhat puzzling. However, the signs of the coefficients are all in predicted direction. Also, note that we do not estimate the effect on 'overconsumption' of unhealthy foods (e.g. candies and soft drinks), which is definitely an important risk factor of overweight. However, we are not able to address this due to data limitation; therefore it should to be addressed in future research.

To summarise, we find that mothers' schooling may improve children's health behaviour. However the effect is not statistically significant.

7 Concluding remarks

In this study we investigate the causal effect of mothers' schooling on children's weight by exploiting the 1972 education reform in England and Wales, which raised the minimum leaving age from fifteen to sixteen. We use the Health Survey for England (1998-2002). We focus on mothers who were born between September 1952 and August 1962, and left school at age fifteen or sixteen. We conduct a regression-discontinuity analysis (using both parametric and nonparametric approaches). We find that the extra year of schooling induced by the reform significantly reduces sons' BMI Z-scores. We do not find any causal effect for daughters. The result provides a supporting evidence for the general education policies to improve national health. Our result is different from Lindeboom et al. (2009), who find no effect of parental schooling on childrens' weight using samples from children who were born in 1958. Note that our results may not be generalised to the general population because we use the restricted samples and moreover the estimates just represent the weighted local average treatment effects. However, we believe our findings are still interesting because less educated individuals (those who left school at the minimum school leaving age) are often targeted in relevant policy discussions.

We fail to find that mothers' schooling increases children's fruit and vegetable consumption and daily physical activity. This result is somewhat puzzling because weight is basically determined by calorie intake and expenditure. Future work could focus on other important lifestyle variables such as snacking, soft drink consumption, and television watching. While these variables are closely related to overweight among children (Ludwig et al., 2001; Andersen et al., 1998), the causal relationship with mothers' schooling has not been examined. Future study should investigate such specific channels throught which the effect of schooling transmits intergenerationally.

References

[1] Anderberg, D. Zhu, Y. 2010 The effect of education on marital status and partner characteristics: evidence from the UK. CESifo Working Paper 3104.

[2] Andersen, R., Crespo, C., Bartlett, S., Cheskin, L., Pratt, M. 1998 Relationship of Physical Activity and Television Watching With Body Weight and Level of Fatness Among Children. Journal of the American Medical Association 279(12) 938-942.

[3] Black, S. Devereux, P Salvanes, K. 2005 Why the apple doesn't fall far: understanding intergenerational transmission of human capital. American Economic Review 95(1) 437-449.

[4] Black, S. Devereux, P. Salvanes, K. 2008 Staying in the classroom and out of the maternity ward? The effect of compulsory schooling laws on teenage births.

The Economic Journal 118(530) 1025–1054.

[5] Brunello, G., Michaud, P-C., Sanz-de-Galdeano, A. 2009 Years of schooling, human capital and the body mass index of European females. IZA Discussion paper 4667.

[6] Cameron, C. Trivedi, P. 2005 Microeconometrics: methods and applications. Cambridge University Press.

[7] Case, A. Lubotsky, D. Paxson, C. 2002 Economic status and health in childhood: the origins of the gradient. American Economic Review 92(5) 1308-1334.

[8] Chen, Y. Li, H. 2009 Mother's education and child health: is there a nurturing effect? Journal of Health Economics 28 413-426.

[9] Chou, S. Liu, J. Grossman, M. Joyce, T. 2010 Parental education and child health: evidence from a natural experiment in Taiwan. American Economic Journal: Applied Economics 2(1) 33-61.

[10] Clark, D. Royer, H. 2010 The effect of education on adult health and mortality: evidence from Britain. NBER Working Paper 16013.

[11] Cole, T. Bellizzi, M. Flegal, K. Dietz, W. 2000 Establishing a standard definition for child overweight and obesity worldwide: international survey. British Medical Journal 320 1240–1243.

[12] Cole, T. Freeman, J. Preece, M. 1995 Body mass index reference curves for the UK, 1990. Archives of Disease in Childhood 73 25-29.

[13] Currie, A. Shields, M. Wheatley Price, S. 2007 The child health/family income gradient: evidence from England. Journal of Health Economics 26 213-232.

[14] Currie, J. 2009 Healthy, wealthy, and wise: socioeconomic status, poor health in childhood, and human capital development. Journal of Economic Literature 47(1) 87-122.

[15] Currie, J. Moretti, E. 2003 Mother's education and the intergenerational transmission of human capital: evidence from college openings. Quarterly Journal of Economics 118(4) 1495-1532.

[16] Currie, J. Stabile, M. 2003 Socioeconomic Status and Child Health: Why Is the Relationship Stronger for Older Children? American Economic Review 93(5)1813-1823.

[17] Del Bono, E. Galindo-Rueda, F. 2007 The long term impacts of compulsory schooling: evidence from a natural experiment in school leaving dates. Discussion Paper CEEDP 74 Centre for the Economics of Education, London School of Economics and Political Science.

[18] Ding, W. Lehrera, S. Rosenquist, J. Audrain-McGovern, J. 2006 The impact of poor health on academic performance: New evidence using genetic markers. Journal of Health Economics 28(3) 578-597.

[19] Doyle, O. Harmon, C. Walker, I. 2007 The impact of parental income and education on child health: further evidence from England. Warwick Economic Research Papers No.788.

[20] Ebbeling, C. Pawlak, D. Ludwig, D. 2002 Childhood obesity: public-health crisis, common sense cure. The Lancet 360(9331) 473-482.

[21] Hahn J. Todd, E. van der Klaauw, W. 2001 Identification and estimation of treatment effects with a regression-discontinuity design. Econometrica 69 201–209.

[22] Imbens, G. Lemieux, T. 2008 Regression discontinuity designs: a guide to practice. Journal of Econometrics 142 615-635.

[23] Lee, D. Card,D. 2008 Regression discontinuity inference with specification error. Journal of Econometrics 142 655-674.

[24] Lee, D. Lemieux, T. 2010 Regression discontinuity designs in economics Journal of Economic Literature 48(2) 281–355.

[25] Lindeboom, M. Llena-Nazal, A. van der Klaauw, B. 2009 Parental education and child health: Evidence from a schooling reform. Journal of Health Economics 28 109-131.

[26] Lleras-Muney, A. 2005 The relationship between education and adult mortality in the United States. Review of Economic Studies 72(1) 189-221.

[27] Lobstein, T. Baur, L. Uauy, R. 2004 Obesity in children and young people: a crisis in public health. Obesity Reviews 5 4-85.

[28] Ludwig, J. Miller, D. 2007 Does head start improve children's life chances? Evidence from a regression discontinuity design. Quarterly Journal of Economics 122 159-208.

[29] Ludwig, D., Peterson, K., Gortmaker, S. 2001 Relation between consumption of sugar-sweetened drinks and childhood obesity: a prospective, observational analysis. The Lancet 357(9255) 17 505-508.

[30] McCrary, J. Royer, H. 2006 The effect of female education on fertility and infant health: evidence from school entry policies using exact date of birth. NBER Working paper 12329.

[31] Powell, L. Bao, Y. 2009 Food prices, access to food outlets and child weight. Economics & Human Biology 7(1) 64-72.

[32] Porter, J. 2003 Estimation on the regression discontinuity model. Unpublished manuscript Harvard University.

[33] Rona, R. Chinn, S. 2001 Prevalence and trends in overweight and obesity in three cross sectional studies of British children, 1974-94. British Medical Journal 322 24-26.

[34] Staiger, D. Stock, J. 1997 Instrumental variables regression with weak instruments. Econometrica 65 557-586.

[35] Stifel, D. Averett, S. 2009 Childhood overweight in the United States: a quantile regression approach. Economics and Human Biology 7 387-397.

[36] van Kippersluis, H. O'Donnell, O. van Doorslaer, E. 2009 Long run returns to education: does schooling lead to an extended old age? Tinbergen Institute Discussion Paper TI 2009-037/3.

[37] Webbink, D. Martin, G. Visscher, M. 2010 Does education reduce the probability of being overweight? Journal of Health Economics 29 29-38.

Table 1: Descriptive statistics of children's weight

A: Full sample				
	Son		**Daughter**	
	Pre-reform	Post-reform	Pre-reform	Post-reform
BMI Z-score	0.597 (1.147)	0.579 (1.118)	0.512 (1.107)	0.545 (1.100)
Overweight rate	25.41%	23.40%	24.96%	28.52%
Observations	669	1222	669	1143
B: Restricted sample (mother left school at 15 or 16)				
	Son		**Daughter**	
	Pre-reform	Post-reform	Pre-reform	Post-reform
BMI Z-score	0.652 (1.141)	0.557 (1.114)	0.572 (1.172)	0.613 (1.088)
Overweight rate	27.51%	23.33%	27.07%	30.03%
Observations	309	660	314	596

Notes: Standard deviations are presented in parenthesis. In Pre-reform, mothers were born between September 1952 and August 1957. In Post-reform, mothers were born between September 1957 and August 1962.

Table 2: Effect of the 1972 reform on mother's school leaving age

	(1)	(2)	(3)	(4)
A: September 1952 to August 1962 cohort (60 months on either side of the cut-off point, N=1879)				
Reform	0.365***	0.378***	0.442***	0.460***
	(0.024)	(0.031)	(0.018)	(0.016)
Polynomial speficication	Linear	Linear	Quadratic	Quadratic
Additional control	No	Yes	No	Yes
AIC	1968.54	1981.70	1970.81	1983.19
G-statistics	0.675	0.839	0.639	0.719
B: Robustness check				
50 months on either side of the cut-off point (N=1583)				
Reform	0.385***	0.411***	0.443***	0.476***
	(0.024)	(0.030)	(0.024)	(0.033)
40 months on either side of the cut-off point (N=1252)				
Reform	0.409***	0.414***	0.398***	0.436***
	(0.016)	(0.031)	(0.022)	(0.023)

Notes: Robust standard errors in parentheses. *** significant at 1% level, **significant at 5% level, * significant at 10% level. All models include polynomials in birth cohort (in ten months) fully interacted with the reform dummy. Additional control includes age of mother, squared age of mother, age of children, squared age of children, dummies for region of residence, and dummies for survey years.

Table 3: Reduced-form effect of the 1972 reform on children's BMI Z-score

	(1)	(2)	(3)	(4)
A: Son (N=969)				
Reform	-0.247*	-0.201*	-0.453***	-0.411***
	(0.136)	(0.104)	(0.103)	(0.108)
Polynomial speficication	Linear	Linear	Quadratic	Quadratic
Additional control	No	Yes	No	Yes
AIC	2977.43	2959.89	2976.82	2957.19
G-statistics	0.878	0.662	0.469	0.463
B: Daughter (N=910)				
Reform	-0.026	0.020	-0.288	-0.347
	(0.215)	(0.146)	(0.190)	(0.190)
Polynomial speficication	Linear	Linear	Quadratic	Quadratic
Additional control	No	Yes	No	Yes
AIC	2790.38	2777.22	2780.23	2763.36
G-statistics	2.456	2.310	1.120	1.250

Notes: Robust standard errors in parentheses. *** significant at 1% level, **significant at 5% level, * significant at 10% level. All models include polynomials in birth cohort (in ten months) fully interacted with the reform dummy. Additional control includes age of mother, squared age of mother, age of children, squared age of children, dummies for region of residence, and dummies for survey years.

Table 4: RDD estimate of the effect of mother's schooling on children's BMI Z-score

	Polynomial control approach				Local linear regression approach		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
A: Son							
Reform	-0.734*	-0.552**	-0.920***	-0.809***	-0.981*	-1.230**	-1.087**
	(0.375)	(0.205)	(0.154)	(0.160)	(0.555)	(0.555)	(0.542)
Polynomial speficication	Linear	Linear	Quadratic	Quadratic	.	.	.
Additional control	No	Yes	No	Yes	No	No	No
Bandwidths selection	.	.	.	.	Optimal	Narrow	Wide
N	969	969	969	969	543	373	698
B: Daughter							
Reform	-0.067	0.052	-0.751	-0.858	-0.050	-0.240	-0.167
	(0.549)	(0.382)	(0.527)	(0.509)	(0.394)	(0.348)	(0.375)
Polynomial order	Linear	Linear	Quadratic	Quadratic	.	.	.
Additional control	No	Yes	No	Yes	No	No	No
Bandwidths selection	.	.	.	.	Optimal	Narrow	Wide
N	910	910	910	910	859	710	996

Notes: Robust standard errors in parentheses. *** significant at 1% level, **significant at 5% level, * significant at 10% level. For results from the polynomial control approach, all models include polynomials in birth cohort (in ten months) fully interacted with the reform dummy. Additional control includes age of mother, squared age of mother, age of children, squared age of children, dummies for region of residence, and dummies for survey years. For results from the local linear regression approach, we select the optimal bandwidths by the cross-validation procedure suggested by Imbens and Lemieux (2008). We also present the results using a narrower bandwidth (-10 months) and a wider bandwidth (+10 months) are also reported. We use the rectangular kernel in local linear regression by following Imbens and Lemieux (2008) and Lee and Lemieux (2010).

Table 5: RDD estimate of the effect of mother's schooling on children's health-related behaviours

A: Children's fruit & vegetable consumption				
	Son (N=354)		**Daughter (N=317)**	
	(1)	(2)	(3)	(4)
Reform	0.747	3.000	0.191	0.392
	(2.487)	(2.991)	(0.522)	(0.844)
Polynomial order	Linear	Quadratic	Linear	Quadratic
Additional control	No	No	No	No
Mean of dependent variable	2.65		2.46	
B: Children's sports (binary)				
	Son (N=537)		**Daughter (N=522)**	
	(5)	(6)	(7)	(8)
Reform	0.112	0.014	0.119	0.103
	(0.147)	(0.097)	(0.120)	(0.132)
Polynomial order	Linear	Quadratic	Linear	Quadratic
Additional control	No	No	No	No
Mean of dependent variable	0.67		0.57	

Notes: Robust standard errors are presented in parentheses. *** significant at 1% level, ** significant at 5% level, * significant at 10% level. All models include polynomials in birth cohort (in ten months) fully interacted with the reform dummy. Mean of dependent variable gives the mean of the outcome variable for the last pre-refrom cohort.

Figure 1: Effect of the 1972 reform on mother's school leaving age (full sample)

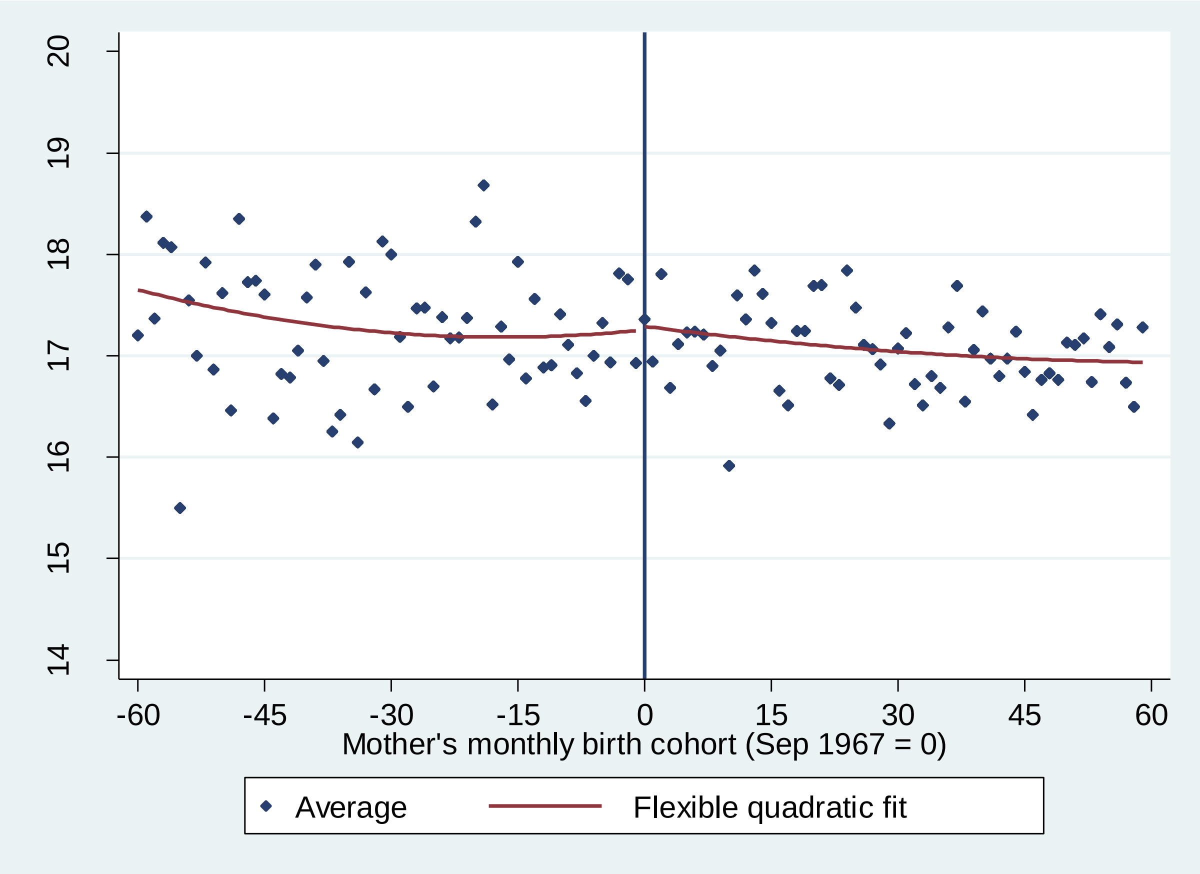


Figure 2: Effect of the 1972 reform on mother's school leaving age (restricted sample)

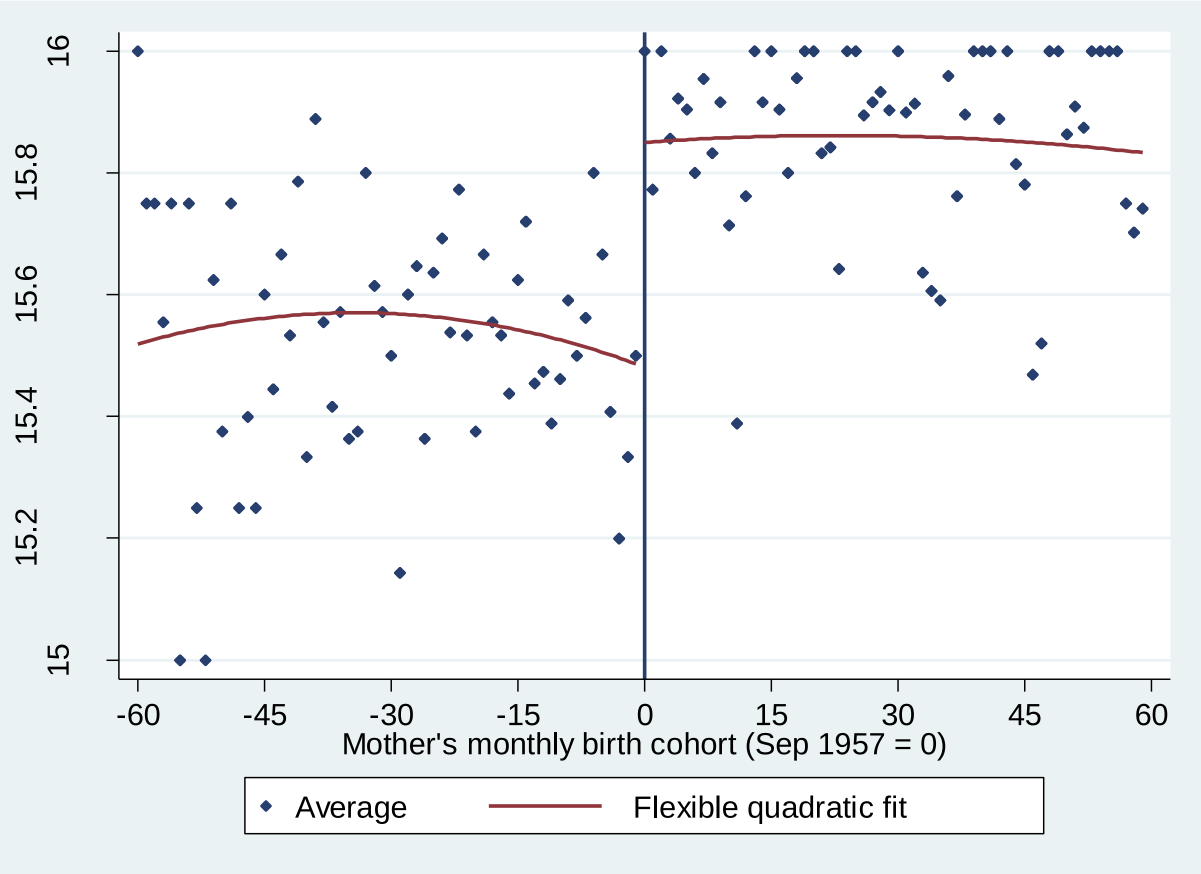


Figure 3: Effect of the 1972 reform on son's BMI Z-score

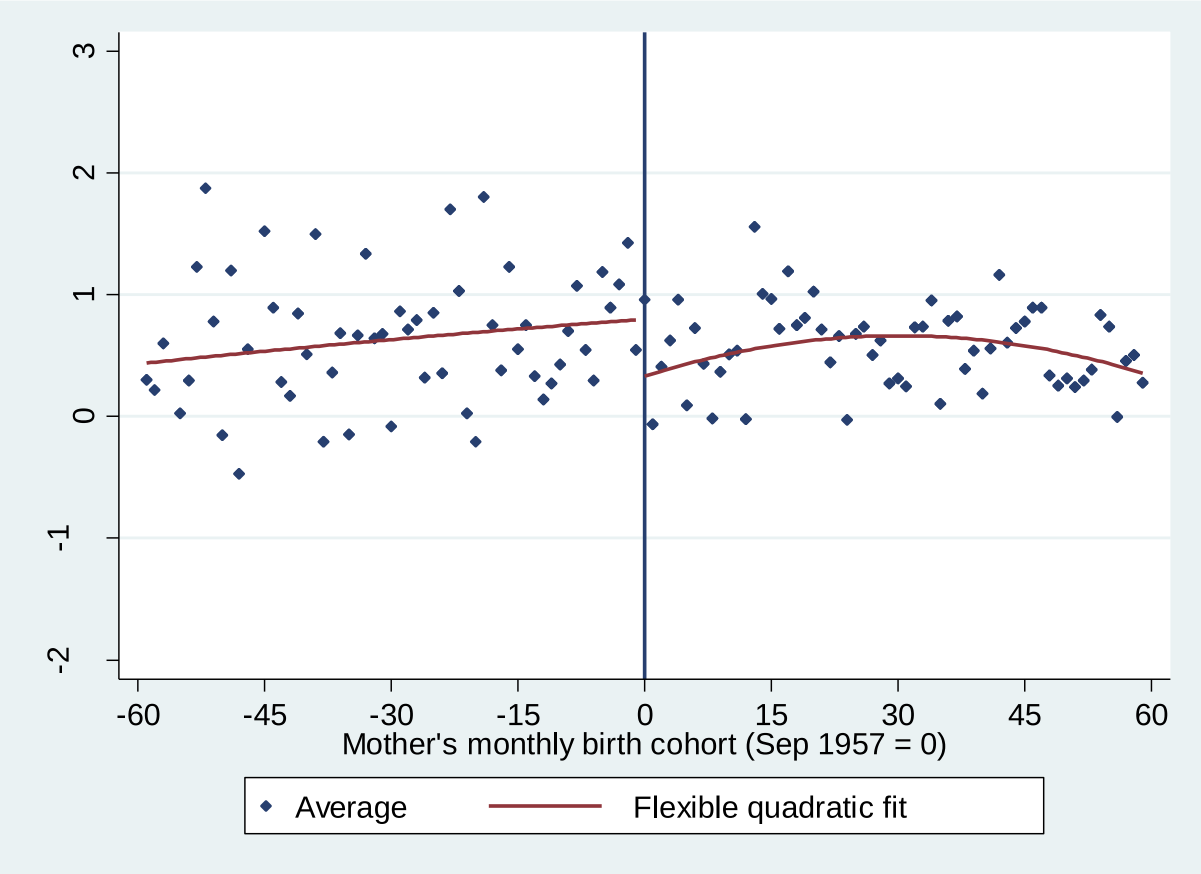
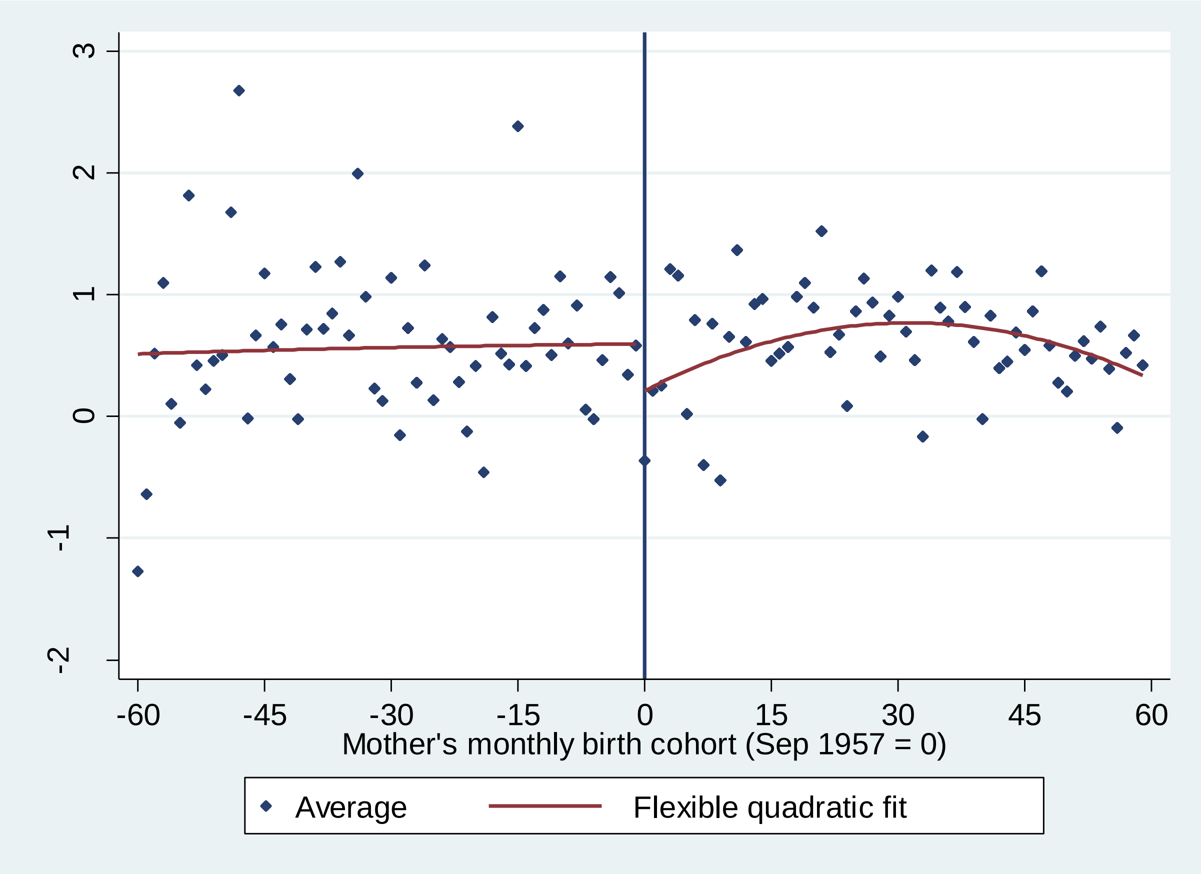


Figure 4: Effect of the 1972 reform on daughter's BMI Z-score



Appendix: Descriptive statistics of other variables

	Mean	Standard deviation	Observations
Mother's characteristics			
School leaving age (for full sample)	17.12	1.91	3605
School leaving age (for restricted sample)	15.75	0.43	1879
Age at the survey (for full sample)	43.9	2.72	3605
Age at the survey (for restricted sample)	43.66	2.66	1879
Son's characteristics			
BMI Z-score	0.59	1.12	969
Overweight rate	24.7%	.	969
Fruit & vegetable consumption (in portion)	2.38	1.99	354
Sports (previous week; binary)	67.4%	.	537
Age at the survey (for full sample)	10.74	3.03	1836
Age at the survey (for restricted sample)	11.00	2.96	969
Daughter's characteristics			
BMI Z-score	0.60	1.12	910
Overweight rate	29.0%	.	910
Fruit & vegetable consumption (in portion)	2.55	2.18	317
Sports (previous week; binary)	60.5%	.	522
Age at the survey (for full sample)	10.70	3.04	1753
Age at the survey (for restricted sample)	11.00	3.04	910

Notes: Summary statistics from the restricted sample are presented (otherwise indicated). Restricted sample includes only mothers who left school at age fifteen or sixteen. Descriptive statistics for region of residence and survey year dummies are omitted.