

NBER WORKING PAPER SERIES

IS THE 'QUARTER OF BIRTH' ENDOGENOUS? EVIDENCE FROM ONE MILLION
SIBLINGS IN TAIWAN

Elliott Fan
Jin-Tan Liu
Yen-Chien Chen

Working Paper 20444
<http://www.nber.org/papers/w20444>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
August 2014

The authors thank to the Ministries of Education and Interior Affairs of Taiwan for providing administrative data. We are grateful for helpful discussions with Stacey Chen, Kamhon Kan, and seminar participants at the Institute of Economics, Academia Sinica, Taipei, Taiwan. We acknowledge Pei-Chuan Ho and Szu-Yu Zoe Kao for their capable research assistance. All errors are our responsibility. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2014 by Elliott Fan, Jin-Tan Liu, and Yen-Chien Chen. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Is the ‘Quarter of Birth’ Endogenous? Evidence From One Million Siblings in Taiwan
Elliott Fan, Jin-Tan Liu, and Yen-Chien Chen
NBER Working Paper No. 20444
August 2014
JEL No. C10,J11,J13

ABSTRACT

Recent studies based on US data have provided evidence to suggest that the ‘quarter of birth’ (QOB) may be endogenous and that the use of QOB as an instrumental variable will consequently produce inconsistent estimates (see Buckles and Hungerman, 2013). Such potential endogeneity is addressed in this study by estimating the effects of QOB on university attendance using a Taiwanese dataset on approximately one million siblings. Our estimations are mainly reliant upon the strength of the family fixed-effects model, a regression discontinuity design and a simulation procedure. Our results, in sharp contrast to the US findings, suggest that family background characteristics can explain very little of the relationship between QOB and the probability of university attendance at the age of 18. The disparity between the US and Taiwanese findings may be due to high-‘socioeconomic status’ (SES) women in the US disproportionately planning births away from the winter months, as suggested by Buckles and Hungerman (2013), whereas the seasonality of births is virtually identical for low- and high-SES mothers in Taiwan. Our findings imply that the endogeneity of QOB is of less concern in the case of Taiwan, perhaps due to the milder winter climate.

Elliott Fan
National Taiwan University
No. 1, Sec. 4, Roosevelt Road
Taipei, 106 Taiwan
elliottfan@ntu.edu.tw

Yen-Chien Chen
National Chinan University
Taiwan
Yenchien@ncnu.edu.tw

Jin-Tan Liu
Department of Economics
National Taiwan University
21 Hsu-Chow Road
Taipei (100), TAIWAN
and NBER
liujt@ntu.edu.tw

1. Introduction

The two decades that have passed since the pioneering work of Angrist and Krueger (1991, 1992) have witnessed considerable development of research into the role of ‘quarter of birth’ (QOB) and its predictive capability on the outcomes of individuals.¹ The key assumption underpinning the validity of the empirical strategies employed within the related studies is that QOB is determined exogenously. The legitimacy of this assumption, however, has been challenged by evidence suggesting that parents may be manipulating the timing of their births according to their preferences for certain birth months or birth seasons.²

As the most recent, and most compelling, example of such preferences, Buckle and Hungerman (2013) found that in the US, the characteristics of women giving birth in the winter months differ significantly from those of women giving birth in other seasons, ultimately leading to a strong correlation between the characteristics of the mothers and the QOBs of their children. They also found that the use of birth season as an instrumental variable yields biased estimates of the returns to education in terms of earnings.³ More interestingly, they found that the seasonality of maternal characteristics was driven by high-‘socioeconomic status’ (SES) mothers with the tendency to avoid winter births, highlighting the potential role of weather in determining the endogeneity of QOB.

It may therefore be of interest to examine the extent to which the US experience can

¹ One particular string of the related literature makes use of QOB as an ‘instrumental variable’ (IV) in an attempt to eliminate potential biases when estimating the returns to education, in terms of earnings or other outcome variables. For recent examples, refer to Lefgren and McIntyre (2006), Chernozhukov and Hansen (2006) Leigh and Ryan (2008), Maurin and Moschion (2009), Arkes (2010), Lee and Orazem (2010) and Robertson (2011). Another line taken in the recent studies is to investigate the ways in which QOB directly affects the outcomes of individuals from perspectives of either educational achievement (Bedard and Dhuey, 2006; and Claire, Dearden and Meghir, 2010) or health (Weber, Prossinger and Seidler, 1998; van Hanswijck de Jonge, Waller and Stetter, 2003; Costa and Lahey, 2005; and Lokshin and Radyakin, 2012), both of which are primary determinants of earnings.

² The validity of using QOB as an IV has also been challenged from two other perspectives. Firstly, significant inconsistency may occur in the IV estimates when there is a weak correlation between QOB and educational outcomes, an argument first put forward by Bound, Jaeger and Baker (1995) following their re-examination of the results of Angrist and Krueger (1992). Secondly, it was recently pointed out by Barua and Lang (2012) that the use of either QOB or legal entry age as an IV also violates the monotonicity assumption, leading to inconsistent estimates of the ‘local average treatment effect’ (LATE). As regards the latter challenge, we present evidence later in this study to suggest that the monotonicity assumption is indeed essentially satisfied in the case of Taiwan; this is mainly because, although there are a number of early school starters (children entering school before attaining the age of six), it is quite rare to find late school starters (age-eligible children delaying school entry) due to ‘redshirting’.

³ These findings clearly point to the potential endogeneity of QOB, echoing the study of Bound and Jaeger (2000) who found that QOB failed to satisfy the requirement of exclusion restrictions when used to instrument education in the estimation of returns, in terms of earnings.

be extrapolated to other countries, and whether the endogeneity of QOB is a common threat to the related studies in which QOB is used as a source of exogenous variations in education. In this study, we aim to contribute valuable evidence to the extant literature by tackling these questions based upon an exploration of the case of Taiwan, where the weather is generally mild in the winter, and thus, the primary ‘suspect’ plaguing QOB evaluation in the US is essentially rendered irrelevant.

We develop multiple strategies to carefully evaluate the potential endogeneity of ‘month of birth’ (MOB), using a unique administrative dataset comprising of around a million siblings (co-residing or otherwise) in Taiwan. Our empirical work focuses on the estimation of the MOB effects on university attendance at the age of 18, with special attention being paid to the role of family background characteristics in determining the magnitude of the estimated MOB effects. We then compare our results to the US-based findings of Buckles and Hungerman (2013) and analyze the parental preferences for winter births in both countries.

In addition to addressing the endogeneity of MOB, we also contribute to the literature in other ways. Firstly, as pointed out by Lokshin and Radyakin (2012), the existing studies carrying out investigations into the QOB effects have thus far been disproportionately clustered in high-income countries. By adopting a dataset on births in Taiwan, we provide evidence on the QOB effects, particularly the long term impacts of QOB, in the context of developing countries.

Secondly, we also estimate the MOB effects on the likelihood of attending public universities in Taiwan, since they are more selective than private universities. Our results should help to advance our understanding not only of the ways in which the MOB affects the probability of university attendance, but also of the effects on the quality of the university attended.

Three empirical strategies are used in our attempt to determine whether the QOB effects obtained from the OLS estimations may be biased as a result of any unobserved family characteristics. Using the sibling data, our first strategy relies upon the application of a family fixed-effects model.⁴ We then compare the fixed-effect (FE) estimates of the QOB effects to

⁴ Lokshin and Radyakin (2012) examined the ways in which the health of children was associated with their MOB, based upon a fixed-effects model with a relatively small sample of siblings. Their FE estimates of the MOB effects on the health of young children are twice the size of the corresponding OLS estimates. It should, however,

the corresponding OLS estimates without controlling for family background variables, in order to assess any potential bias of the OLS estimates. To the extent that family characteristics are invariant across siblings, the FE model suppresses the correlation between QOB and these characteristics, thereby producing more credible estimates than the OLS alternative. If family background characteristics do not have any relevance to the decisions made by parents with regard to QOB, then no significant FE-OLS disparity will be observable.

Our second strategy makes use of a regression discontinuity design (RDD), which exploits a single starting date for schools in Taiwan; that is, the first of September. Under this setting, the cut-off point is set as the first of September along with the ‘forcing’ variable, which is defined as the dates of birth of individuals. The RDD therefore compares the likelihood of university attendance at the age of 18 for individuals born just before, and just after, the first of September. The hypothesis supporting this comparison is that dates of birth occurring around the first of September are essentially randomly determined, given that parents are unable, or had no specific intention, to exercise precise control over the birth timing around the cut-off date.

Our third strategy involves the use of our sibling sample to carry out a simulation which enables us to investigate whether the dates of birth of the sample are centralized around certain months or seasons. Such centralization, if present, would indicate parental selection of the birth seasons, or the birth months, of the children according to the preferences of the parents. In this simulation, the actual date of birth of each individual is replaced by a simulated date of birth, randomly drawn (without replacement) from the empirical distribution of the dates of birth contained within the original sibling sample. We then calibrate and compare the average gap between the dates of birth of siblings using both the simulated and actual dates of birth. Given that the former will be free from any form of parental selection, it should be greater than the latter if centralization exists. Therefore, such inequality lays the foundation for the key test of our hypothesis in this study.

Our FE estimates of the MOB effect are surprisingly large: the probability of university attendance at the age of 18 for males born in September is about 34.4 per cent higher than the probability for males born in August, whilst the margin for females is around 25.3 per cent. Both are very close to the corresponding RDD estimates, thereby raising our

be noted that, as argued by Lokshin and Radyakin (2012), the siblings sample used in their study may be subject to selection biases and measurement errors, which could explain the differences between the OLS and FE estimates.

level of confidence in the reliability of the results. More importantly, the disparity that exists between the FE estimation and the OLS alternative (with no controls in place for the family background variables) is negligible for males, although the margin is somewhat larger for females. Our simulation results also show that any clustering of the dates of birth amongst our sample appears to be only in line with the general seasonal pattern of births, and not indicative of any significant degree of parental selection. Our results are not consistent with the suggestion that the OLS estimates are confounded by the influences of the family background characteristics, thereby presenting a scenario which stands in sharp contrast to that portrayed by Buckles and Hungerman (2013) based on the US births.

Further examination of this disparity between the US and Taiwan cases reveals that the seasonality of births is virtually identical for both high- and low-SES mothers in Taiwan. In the case of the US, however, Buckles and Hungerman (2013) found that high-SES women disproportionately planned births away from the winter months, thereby leading to different seasonal patterns of births between high-SES and low-SES women, and thus, a strong correlation between QOB and family characteristics. Perhaps the Taiwan case is completely different from the US case because when contemplating their birth plans, parents in Taiwan do not need to avoid the winter months, since the winter weather in Taiwan is generally very mild. More evidence is, however, required to shed light on this supposition before it can be further accepted.

2. Background

2.1. Education system in Taiwan

The educational system in Taiwan is largely similar to that of many Western countries, with students beginning their education at primary school, which requires a total of six years to complete, and then supplementing this with three years in a junior high school. Upon completion of junior high school, students have the option to continue studying for three years in an academic senior high school or to enter a vocational secondary school for two to five years. Normally, those who complete the academic senior high school program participate in the College Entrance Test (CET) with the aim of securing their undergraduate admissions.

In July of each year, on completion of the CET, each applicant submits an application

form containing their choices of department across all of the universities to which they would like to apply. The departments listed in the application form are ranked in their order of preference by the applicant, with a maximum number of departments to which any individual may apply.⁵

The higher education system in Taiwan can be categorized into two tiers, the more prestigious public universities and the less selective private universities, with public universities being the primary choice for most students.⁶ Final admission is determined by a student's test score and the rankings of the schools listed in the application form. An applicant is subsequently granted admission to only one department, and must then decide whether or not to accept admission to that department. Regardless of the admission outcome, each student is free to take the CET again and enter the new round of admission applications in the following year.

The single school starting date has been fixed as the first of September in Taiwan since the mid-twentieth century. All children are permitted to enter elementary school, provided that they have attained the age of six years by the first of September of the relevant year. As in the case of many other countries, not all parents in Taiwan strictly comply with the legal requirements relating to school starting age. However, what is unique about the case of Taiwan is that, whilst there are a number of early starters (children entering elementary school prior to reaching the age of six), it is quite a rarity to find cases of late starters (age-eligible children delaying school entry).

According to the information collected in the Taiwan *Survey of Family Income and Expenditure* (SFIE), which relates to children born between 1993 and 2000, the proportion of early starters during that period was around 6.6 per cent, whilst late starters (including those who never started school) accounted for only 2.8 per cent.⁷ Although our dataset provides no information on the actual school starting age of each individual, in order to

⁵ The maximum number varied across years. It was 99 in 1994, 66 during 1995 and 2001, 80 during 2002 and 2003, 90 in 2004, and 100 afterwards.

⁶ Public universities receive considerably greater government financial aid, as compared to private universities, regardless of whether this is assessed in terms of the total subsidy or the subsidy per student. As a result, tuition fees for an academic year range between NT\$26,680 and NT\$35,940 for public universities, whereas for private universities, the fee are more than twice as high, ranging from NT\$67,680 to NT\$93,860.

⁷ Due to a lack of information in the SFIE data, we are unable to observe whether these late starters are enrolled in an elementary school at a later date. Given the possibility that some children are unable to go to school due to illness or other causes, it is highly likely that the actual proportion of late starters will be lower than the 2.8 per cent stated here.

ensure the thoroughness of the analysis in our study, we carry out an evaluation of the effects of having these non-compliant children on our estimates of the MOB effects in Section 5.4.

2.2. Seasonality of births and maternal characteristics

In this section, we focus on documenting the seasonal pattern of births and compare this with the seasonal pattern of the characteristics of the mothers and children observed in our sample. Our comparison is quite intriguing, essentially because Buckles and Hungerman (2013) made similar comparisons to identify the correlation between QOB and parental characteristics. The maternal characteristics selected for this study correspond largely to those based upon US data presented in Buckles and Hungerman (2013), so we are able to make a preliminary comparison of the two countries with respect to the seasonality of maternal characteristics.

A strong seasonal pattern in the characteristics of the mothers would imply that those with different characteristics may hold different preferences towards or against certain seasons in which they would like to give births. It is, however, important to note that such seasonality does not, by itself, imply that the OLS estimates of the MOB effects are plagued by omitted variables bias (OVB) when family background factors in the estimation are not fully controlled for. Such bias occurs when the factors driving the seasonal cycle of births also drive the educational outcomes of the children, thereby leading to a correlation between MOB and the errors. This is exactly the case in the US, where a strong correlation is indicated between the seasonality of births and the seasonality of maternal characteristics.⁸

Indeed, as documented by Buckles and Hungerman (2013), the number of births in the US is particularly small in January. At the same time, children born in January are disproportionately likely to be born to women who are teenagers, unmarried and lacking a high school degree. The two authors argued that this was likely to be caused by high-SES mothers avoiding the timing of their births in the winter months, which is perhaps a factor leading to the biases in the OLS estimates of the QOB effects when maternal characteristics are not appropriately controlled for.

The scenario in the case of Taiwan is, however, markedly different. We examine the seasonality of these variables using samples drawn primarily from the birth registry records

⁸ Using physics-based terminology, the seasonality pattern of births is in phase (or in anti-phase) with the patterns of maternal characteristics.

in Taiwan, since these contain information on every single birth that occurred between 1978 and 1984. The results are illustrated in Figure 1, with the seasonal pattern of births being shown in graph (A), from which we can see that like many other countries, the number of births exhibits a substantial and persistent pattern of seasonality (Lam and Miron, 1991).⁹

For each of the years examined, October is the month with the highest number of births, whilst April to be the month with the lowest number.¹⁰ This pattern of seasonality is inconsistent with the pattern found in the US, where the trough is located in December or January and the peak in September. Furthermore, the two countries also differ in terms of the magnitude of the seasonality. In Taiwan, the number of births in the peak month is around 30 per cent higher than the number in the trough month of the same year, whereas in the US, the difference is only about 15 per cent.

The proportion of the mothers in the sample with high school or higher education is illustrated in graph (B) of Figure 1, where a marked growth trend is discernible in the proportion of mothers completing high school education. Nevertheless, the curve does not exhibit any clear pattern of seasonality, and neither the peaks nor the troughs are located in October or April, where both the peaks and troughs of the births routinely occur.

Thus, unlike the US case, where winter births are noticeably fewer and the educational level of those mothers giving birth in the winter is markedly lower, no such correlation can be uncovered in graphs (A) and (B). Further, the seasonal pattern of other maternal characteristics also differs from that of births. Similar to the bumpy curve in graph (B), an unclear pattern of seasonality is shown in graph (C) with regard to the proportion of children born to unmarried mothers.

Finally, we add graph (D) to Figure 1 to examine whether there is any seasonality in the proportion of low birth-weight babies.¹¹ Indeed, the seasonality shown in graph (D) clearly suggests that babies born in the winter months are more likely to fall into the category of low birth-weight babies, as compared to the summer born. But, again, the locations of the peaks and troughs do not coincide with those for the numbers of births. Thus, the graphs presented in Figure 1 do not appear to indicate any correlation between the MOB of the child

⁹ The exclusion rules for our estimation samples are detailed in Section 4.

¹⁰ In fact, this seasonality pattern remains unchanged throughout all of the years up to the 2000s, as revealed in more recent data.

¹¹ A newborn baby is categorized as a low birth-weight baby if its body weight at birth is below 2,500 grams.

and the characteristics of the mother. We go on, later in this study, to verify this graphical implication using econometric strategies.

3. Estimation strategies

It is extremely important that our investigation of the role of MOB on university attendance does not introduce any implications of general equilibrium effects, such as the effects from either expediting or postponing the school starting age for the entire cohort of children as a result of policy changes. In order to shed some light on these general equilibrium effects, there is a requirement to distinguish between the ‘absolute age effect’ and the ‘relative age effect’ of starting school.

The absolute age effect refers to the advantage for children starting school at an older age, whilst the relative age effect concerns the advantage for children being older than their peers. Of these, although the relative age effect is more relevant to individual decisions, it is only the absolute age effect which helps to deal with the question regarding the optimal starting age for the entire population of children. As in all of the prior studies attempting to estimate the MOB effects, we are also unable to discern the absolute and relative age effects; thus, our findings serve only as inferences for individual decisions.

3.1. Sibling strategy

We begin with a parsimonious model describing individual educational outcomes as a linear function of a set of determinants, with the specification being expressed as:

$$Uni_{ij} = \alpha_0 + \sum_{k=1}^{11} (\alpha^k \cdot SM_{ij}^k) + X_{ij}\lambda + Z_j\gamma + \varepsilon_{ij}, \quad (1)$$

where Uni_{ij} is a dummy variable which indicates university attendance in the year that individual i of family j attains the age of 18 ($Uni_{ij}=1$); and SM_{ij}^k refer to ‘school month of birth’ k for individual i in family j . In contrast to calendar months, which of course begin with January in each year, the school month of birth in Taiwan begins with September as the first month of a school year; therefore, $SM_{ij}^1=1$ indicates individuals born in September, $SM_{ij}^2=1$ indicates October-born ones, etc., with August being regarded as the omitted dummy; X_{ij} represents a set of individual demographic variables, including birth weight

and birth year; Z_j is a vector indicating parental characteristics; ε_{ij} is the individual specific error term.

One concern with regard to the estimation of equation (1) lies in the possible correlation between SM_{ij}^k and unobserved family characteristics, all of which are potential components of the error term. In the FE model, the error term can be decomposed into two components, a completely idiosyncratic term, denoted by v_{ij} , and a family-specific term, denoted by u_j , with the latter capturing the genetic and family environmental factors that are fully shared by siblings in the production function of the educational outcome. Thus, equation (1) can be re-written as:

$$Uni_{ij} = \alpha_0 + \sum_{k=1}^{11} (\alpha^k \cdot SM_{ij}^k) + X_{ij}\lambda + Z_j\gamma + (u_j + v_{ij}), \quad (2)$$

where $\varepsilon_{ij} = u_j + v_{ij}$. Estimations of equations (1) and (2) are executed for males and females separately, as it is commonly acknowledged that the MOB effect may differ between genders.¹²

One potential threat to the FE model lies in the possibility that parents may choose to reallocate family resources, through compensatory or reinforcing investments, across siblings in response to children's endowments.¹³ This type of intra-family idiosyncratic elements cannot be eliminated by controlling for family fixed effects as in equation (2). While the existence of these elements may lead to a bias in our FE estimates of α^k , the direction of the bias is still difficult to determine.

If the parents prefer 'equality of outcome' amongst their children, then they will readily redirect resources from their higher achieving children to their lower achieving children. In this case, if the disparity between the performance levels of siblings is at least partially attributable to different birth months, then the FE estimates acquired from the estimation of equation (2) will tend to understate the MOB effects. In contrast, if the parents believe in 'equality of opportunity', then there will be no performance-driven reallocation of resources amongst siblings. Thus, the FE estimates will be unbiased. Furthermore, it is

¹² It has been widely recognized that in early childhood girls mature more quickly than boys. If so, we have a strong reason to believe that younger boys may suffer from double disadvantages when they start school in a mixed-sex class, which is the prevailing type of classes in elementary schools in Taiwan.

¹³ Evidence is quite mixed on how parents differentiate human capital investments across children. A recent study done by Heckman, Yi, and Zhang (2013) suggest that both compensatory and reinforcing investments occur simultaneously, but in different perspectives of human capital, such as education and health.

also feasible that parents will tend to strategically focus their resources on their higher achieving children to achieve the maximum outcome amongst their children, as opposed to the highest average outcome. In that case, the FE estimates will be upwardly biased.

To gauge these potential biases, we run the regressions for two-child families only, which enables us to examine whether the MOB of the second child alters the marginal effects of the MOB of the first child. A significant effect of the MOB of the second child would imply intra-family resource allocation. Details of this examination are laid out in the Appendix.

Another potential threat to the FE strategy is that, in our primary data source, the treatment (that is, the actual school starting age) is unobservable at the individual level. Thus, we can only use the legal age for school entry to carry out our estimations under the assumption that, in compliance with the law, each individual is enrolled in school on the first of September after attaining the age of 6 years. However, as noted in Section 2, it is unrealistic to expect this assumption to hold given that there are children who start school either earlier or later than the legal age requirement. Our estimated MOB effects should therefore be considered more as intent-to-treat (ITT) effects. Furthermore, it would be inappropriate for us to apply the conventional remedy of scaling up the ITT estimates by using the proportion of compliers. This is because those who are non-compliers are not ‘neutral’ in the estimations: they migrate from the treatment group to the control group (as early starters) or vice versa (as late starters), instead of being randomly selected and then excluded from the estimation samples. Later in Section 5.4, we go on to evaluate how our ITT estimate would tend to deviate from the average treatment effect (ATE) due to the existence of these non-compliers.

3.2. Regression discontinuity design

Our second estimation strategy for the MOB effect exploits the merit of a regression discontinuity design (RDD). Under the continuity assumption, our analysis is based on a reduced-form regression as follows:¹⁴

$$Uni_i = \beta_0 + \beta_1 \cdot Eligible_i + \beta_2 \cdot K(SBD_i) + \varepsilon_i, \quad (3)$$

¹⁴ The continuity assumption states that, for the RD design to be legitimate for estimating the treatment effect, it is necessary to assume that all factors other than the treatment and outcome variables are continuous with respect to the forcing variable.

where SBD_i refers to the standardized date of birth for individual i , which is measured by the date of birth minus the cutoff date along the dimension of the calendar;¹⁵ $Eligible_i \equiv 1(0 \leq SBT_i \leq 183)$ is an indicator equaling to one if the individual is born between the first of September and the third of March in the following year. Given the discrete nature of the dates of birth of individuals, the regression approach must be able to extrapolate the precise position of the discontinuity from the data. To achieve this, we adopt the method proposed by Lee and Card (2007) in order to deal with the errors caused by the discrete forcing variable. In specific terms, equation (3) can be expressed in a local linear regression form, as follows:

$$Uni_{ib} = \beta_0 + \beta_1 \cdot Eligible_{ib} + \beta_2 \cdot SBD_{ib} + \beta_3 \cdot SBD_{ib} \times Eligible_{ib} + \varepsilon_{ib}, \quad (4)$$

where the subscript b denotes the date of birth. Following the strategy proposed by Lemieux and Milligan (2008), all of the information at the individual level can be summarized in the day-specific means of the variables, with equation (4) reducing to:

$$Uni_b = \beta_0 + \beta_1 \cdot Eligible_b + \beta_2 \cdot SBD_b + \beta_3 \cdot SBD_b \times Eligible_b + \varepsilon_b. \quad (5)$$

Lemieux and Milligan (2008) argued that the estimates of equation (4) at the individual level would be the same as the weighted estimates of equation (4) at the group (day) level if the number of observations in each date of birth group are used for the estimation weighting, although the standard errors would differ.

We applied the method proposed by Imbens and Kalyanaraman (2012) for our bandwidth selection, with a rectangle kernel being used for weighting.¹⁶ Finally, multifaceted examinations were carried out as tests for the robustness of our estimates to (i) both individual level and group level estimations; (2) various bandwidths covering almost the entire range of the forcing variable values; and (3) various sets of covariates incorporated in the estimation.

4. Data

The data used in this study were obtained by merging two national administrative datasets

¹⁵ It should be noted that the value of SBD_i is limited to the range of $(-183, 183)$. However, this limitation does not have any adverse effects on our estimations, since the selected bandwidths are substantially lower than 183.

¹⁶ The choice of bandwidth in recent empirical studies is based on either cross-validation or ad hoc methods. The conventional cross-validation is to find an optimal bandwidth for fitting a curve over the entire support of the data, but what really matter to the RDD estimate are the boundary values at the cutoff. The algorithm proposed by Imbens and Kalyanaraman (2012) select the optimal bandwidth by minimizing the squared error loss at the cutoff. They thus argue that it has better properties than other bandwidth selection methods.

at the individual level. The first was retrieved from the birth registry records of the entire population in Taiwan, with a total of 7,053,190 births having taken place between 1978 and 1999. These records, which are compiled by the Ministry of Interior Affairs, provide detailed information on individual ID, date of birth, birth order and birth weight of each newborn, as well as the ID, age, education level and residential location of the parents. Thus, we can use the parental ID codes to trace the complete history of births for any woman. Our sibling sample is constructed by linking the birth registry records for those mothers who first gave birth between 1978 and 1984. We then track all of the subsequent births for each mother up to 1999, thereby covering a period of up to 21 years.¹⁷ The resulting data comprise of 1,911,627 births, 987,678 of which were males and 923,949 females.

The second data source is provided by the College Entrance Test (CET) files, covering the years 1996 to 2003, collated by the College and University Entrance Test Center of the Ministry of Education. The CET dataset provides detailed information on the year that an individual participated in the CET, the high school graduation year, the CET score and the university admission results. The outcome variable is measured by a dummy variable which indicates whether an individual is admitted to a general university or a public university in the year that he/she reaches the age of 18 years.

If we are to effectively carry out the separate estimation of males and females using the family fixed effects, it is necessary to construct sibling samples for all brothers and sisters in the original data. Since the CET data window only allows us to observe the university admission results recorded on individuals between 1996 and 2003, our sibling samples can only cover individuals born between September 1978 and August 1984, a total of 6 school years. We then construct a ‘male sibling’ sample, comprising of males for whom there was at least one observation of a brother in the same sample period (regardless of any sisters he may have), and a ‘female sibling’ sample, comprising of females for whom there was at least one observation of a sister in the same sample period (regardless of any brothers she may have).

Both of the samples are subject to right-censoring, essentially because the later in the sample the first child was born, the lower the likelihood of a higher order child, if any, being

¹⁷ In order to ensure that our family FE estimates can be more readily interpreted, a small number of half-siblings and step-siblings of the firstborn child were excluded from our estimation samples.

born during the sample period. However, under the assumption that the censoring criterion is randomly determined, such right-censoring should prove to be harmless. Given that the censoring criterion in this study is determined by a specific calendar period (August 1984), we believe that the assumption of randomness should be well satisfied. The two samples extracted from our original dataset provide us with 452,095 male observations and 424,367 female observations.

Although it was necessary to exclude families with only one boy or one girl from our samples, later, in Section 5.4, we present our tests for the robustness of our estimates to this sample exclusion rule by using these excluded individuals to estimate and compare the MOB effects with the corresponding estimates based on the sibling samples. The results reveal that these two sets of estimates are very similar in magnitude, although, in terms of the percentage change, the effects on the excluded males and females are generally smaller.

Finally, to ensure that the OLS results are comparable with the FE results, our OLS estimations are carried out using the same sibling samples, with the summary statistics for males and females being reported in Table 1. For the 1978-1984 cohort, a higher probability of university attendance is found for females (14 per cent) than males (13 per cent), although the probability of public university attendance is roughly the same for both genders (around 5 per cent). Also similar are the age of the mother and father at the time of the child's birth. As expected, birth weight is higher for males than females and the educational level of the fathers is generally higher than that of the mothers. Finally, females have more siblings than males. This reflects the preferences of parents towards boys, since families with only girls are more likely to have more births in the hope of having a boy.

5. Results

5.1. The FE and OLS results based on the sibling samples

Our ordinary least squared (OLS) and fixed-effects (FE) results based on the male and female sibling samples are respectively presented in Tables 2A and 2B. The significant estimate of the September dummy in the OLS results on the male sample, shown in column (1) of Table 2A, suggests that the likelihood of university attendance at the age of 18 for males born in September is 3.44 percentage points higher than that for males born in August. This margin is quite striking, since it constitutes a 32 per cent increase on the probability of

attendance for August-born males. Moving down the column, the estimates for later months demonstrate a clear declining pattern of the MOB effect, since it falls roughly monotonically from a significant estimate of 3.09 percentage points for October to almost zero (and insignificant) for July.

The corresponding FE results are reported in column (2) of Table 2A, which shows that the estimate for September is 3.68 per cent, only slightly larger than the OLS estimate in the first column. In line with the declining pattern of the OLS estimates, the FE estimates also decline in the later months. In order to provide a better illustration of the FE-OLS disparities, both sets of estimates are plotted in Figure 2.

As shown in graph (A) of Figure 2, for most of the months, the FE estimates remain very close to the corresponding OLS estimates. To formally test the significance of the OLS-FE differences, we apply the Seemingly Unrelated Regression estimation (SURE) by running both the OLS and FE equations simultaneously, allowing for non-zero correlation between the error terms of the two equations.¹⁸

The test results, presented in p -value terms, are reported in column (3) in Table 2A. Throughout the entire column, no OLS-FE pairs suggest any statistically significant difference. The result from the SURE test on the joint equivalence of all eleven OLS-FE pairs, shown at the bottom of column (3), also fails to reject the null hypothesis.

The results on the incidences of public university attendance at the age of 18 present a very similar scenario. As suggested by columns (4) and (5) in Table 2A, for both the OLS and FE estimates, the MOB effect peaks in September, with statistical significance, and declines roughly monotonically in the later months. Furthermore, the OLS-FE difference is generally small, and the p -values from the SURE tests, reported in column (6) in Table 2A, suggest that none of the OLS-FE difference is significantly different from zero. This OLS-FE similarity is corroborated by graph (B) in Figure 2, which shows that the OLS-FE gap is rather limited for most months, and that one may be larger than the other interchangeably.

We now turn to graph (C) in Figure 2 for the estimated MOB effects on the female

¹⁸ The approach of SURE has an advantage over methods, such as the conventional t test, that assume zero cross-equation correlations. Using this SURE model, we can conduct a one-to-one test on the equivalence of the FE estimate and the corresponding OLS estimate, as well as on the equivalence of a set of estimates.

sibling sample. Similar to the results on the male sibling sample, a declining pattern is exhibited by both the OLS and FE estimates, although there is no decline in either the OLS or FE estimates between September and November. However, unlike the minimal OLS-FE differences in the results for the male siblings, it is apparent that for females, the FE estimates are smaller than the corresponding OLS estimates for the entire range of months, with the one exception of September. Despite the graphical OLS-FE differences, the margins are mostly statistically insignificant. Columns (1) and (2) in Table 2B show the complete sets of MOB estimates, with the p -values from the SURE tests being listed in column (3). The results indicate that the OLS-FE difference is only significant at the 1 per cent level for May, and at the 5 per cent level for November and January. The dominance of the OLS estimates over the FE estimates is also clear for the public university attendance of females, as shown in graph (D).

5.2. RDD results

Prior to presenting the RDD results, it is important for us to provide validity tests to back up the RD design adopted in this study. We start with examining whether the means of the observable characteristics are continuous through the cut-off point, since any significant discontinuity in these characteristics indicates that parents with particular characteristics may be able to manipulate the timing of births around the first of September.

For both the male and female sibling samples, the respective mean values of eight observable characteristics against the forcing variable, as measured by the number of days away from the first of September, are illustrated in Figures 3A and 3B, with these figures showing that none of the eight variables exhibits any obvious discontinuity at the cut-off point for either gender, although these are only ‘eyeball’ tests.¹⁹

We go on to adopt a more formal approach involving estimations of local linear regressions based upon the method proposed by Lee and Lemieux (2010). We use each of the observable characteristics as the dependent variable in order to test whether the mean is smooth across the cut-off point. The regression results, as shown by the coefficient estimates of the

¹⁹ The eight variables are father’s age and years of schooling, mother’s age and years of schooling, proportion of first-born children, birth weight, size of siblings (including self) and proportion of individuals residing in urban areas.

dummy T in Table 3, confirm the smoothness as illustrated graphically in Figures 3A and 3B.²⁰

As our second test for the validity of the RD design, we examine whether the density of the observations exhibits any discontinuity at the cut-off point, which, in itself, would represent direct evidence of sorting around the threshold.²¹ To this end, the frequency distributions for males and females across the cut-off date are illustrated in Figure 4, which presents the daily number of births across the first of September (the cut-off point is at zero). For both genders, no discontinuity is exhibited in the frequency distribution at the cut-off point.

Graph (A) in Figure 5 reveals a sharp discontinuity in the probability of university attendance on the first of September for males, with the jump being quite striking, from roughly 10 per cent on the left side of the cut-off point to around 14 per cent on the right side. Such discontinuity remains noticeable for females, as illustrated in graph (B), although the means appear to be bumpier.

The probability of public university attendance is plotted in Figure 6, which shows that for both the male and female sibling samples, the jumps on the first of September are still visible, although less so than in the case of general university attendance.

The respective RDD results from the estimation of equation (5) for males and females are presented in Panels A and B of Table 4. Column (1) in Panel A shows that when the RDD estimation is carried out at the group (day) level, the probability of university attendance at the age of 18 for a male born on the first of September is 3.29 percentage points (or 32.6 per cent) higher than that for a male born one day earlier (on the thirty-first of August). This figure is close to the FE estimate for the September dummy presented in Section 5.1 (3.68 percentage points or 34.4 per cent), which reflects the average difference in the probability of university attendance between males born in September and males born in August.

As expected, the group-level estimate is identical to the individual-level estimate, as shown in column (2) of Panel A, with a slightly reduced standard error. When controlling

²⁰ The only exception is the coefficient estimate of T in the regression of proportion of firstborn children for females. However, the estimate (0.88 percent) is trivial if compared to the sample mean (39.53 percent).

²¹ Given the discrete nature of the forcing variable, we are unable to apply the manipulation test proposed by McCrary (2008), which was designed to examine whether the density of a continuous forcing variable exhibited any discontinuity around the cut-off point.

for no covariates in columns (1) or (2), the estimate changes only slightly after the incorporation of county of residence and year of birth as the control variables (column 3), and after further controlling for a set of family and individual characteristics in the regressions (column 4).

The RDD estimate on the female samples suggests a margin of 3.54 percentage points (or 31.9 per cent) between females born on the first of September and females born one day earlier, as shown in column (4) in Panel B of Table 4. Once again, the estimate is found to be strongly robust to the incorporation of different sets of covariates.

As a cautionary measure, we examine the sensitivity of our RDD estimates to any variations in the bandwidth, from 20 days to 120 days, with the results being illustrated in Figure 7. The dark curve in the figure shows the RDD estimates for males, which is quite stagnant over the entire range of bandwidths. The robustness of the RDD estimates remains strong for females (the light curve), though the curve exhibits a growth from around 0.032 with the bandwidth of 36 to around 0.042 with the bandwidth of 63.

Finally, the results on public university attendance are reported in columns (5) to (8) of Table 4. The respective RDD estimates for males and females, at 0.94 and 0.60 percentage points, are both significant. Also, they are analogous to the corresponding FE estimates shown in Panels A and B of Table 2.

5.3. Why does not family background matter?

The main empirical results of our study suggest that the failure to control for family background variables in the OLS estimations does not result in any serious omitted variables bias (OVB) in the estimates of the MOB effect on the likelihood of university attendance at the age of 18 years. Basic econometrics theory provides two possible reconciliations for this outcome: (i) family background (the omitted variable) does not predict educational attainment (the outcome variable); or (ii) the family background variables are uncorrelated with MOB (the X variables of interest). When any of the two conditions holds, OVB is not a concern. In this section we focus on examinations on the second condition, since the first is unlikely to be authentic.

A straightforward way of exploring the correlation between MOB and the family background variables is to examine whether births to mothers with different socioeconomic

status (SES) levels exhibit different patterns of seasonality. Monthly newborns to mothers completing high school or higher education (hereafter referred to as HS mothers), as compared to those born to mothers with lower education levels (hereafter referred to as non-HS mothers), are illustrated in Figure 8.²²

Here, we use education as an indicator of SES, with each point in the curve representing a standardized value, which is obtained by taking the total number of births in a month and subtracting the average monthly births over the year, and then dividing by the average.²³ We find that for both males, in graph (A), and females, in graph (B), the solid curve (HS mothers) and the bar curve (non-HS mothers) are almost synchronized throughout all months, thereby indicating strongly similar patterns of seasonality for the two groups of mothers.

Our finding provides support for the second condition outlined at the beginning of this sub-section, since such similarity implies that MOB is practically uncorrelated with the educational level of the mothers. We can formally test this similarity by running a regression with a dummy variable for HS mothers on a full set of MOB dummies (August omitted). The results are reported in Table 5. The first column of the table, which presents the results for the female siblings, shows that all of the MOB estimates are negligible and insignificant. For the males, while the coefficient estimates for December, March, and April are statistically significant, the magnitudes of these three estimates (between 0.0079 and 0.0088) remain minimal. Since these regression results corroborate the findings based on the illustration in Figure 8, they indicate only a limited difference in the seasonal pattern of births between HS and non-HS mothers.

Our findings in the present study are clearly at odds with the findings reported on a US sample by Buckles and Hungerman (2013), since they revealed noticeable differences in the seasonal pattern of births between high- and low-SES women.²⁴ The Buckles and Hungerman (2013) study shows that children born in the winter months were disproportionately born to low-SES mothers and that, as compared to those born in other

²² Women graduating from either an academic or vocational high school are considered to have completed their high school education. In our samples, the proportion of HS mothers (educated to high school level or higher) is found to be around 25 per cent.

²³ Doing so eliminates the year effects as well as the mean difference in the number of births between the two groups of mothers, thereby making the two curves comparable.

²⁴ Buckles and Hungerman (2013) used marital status as a proxy for socio-economic status.

months throughout the year. Meanwhile, winter-born individuals tend to end up with lower education and earnings. The two correlations together constitute a typical OVB problem in a regression of education or earnings on QOB variables when no controls are put in place for the family background variables.

This is exactly why Buckles and Hungerman (2013) found that the omission of family characteristics as controls within such a regression led to the overstatement of the QOB effects, and as such, they concluded that the use of QOB as an instrumental variable was inherently problematic. In the case of Taiwan, however, we find only a trivial difference in the seasonal pattern of births between HS mothers and non-HS mothers, which effectively breaks the necessary link causing such OVB. This offers some reconciliation of the similarity found in this study between the OLS estimates and FE estimates obtained in Section 5.1.

5.4. Robustness examinations

Regarding early starters

As noted earlier in Section 2.1, between 1993 and 2000, the proportion of early starters was around 6.6 per cent of the entire cohort, whilst the proportion of late starters (including those who never started school) was, at the very most, 2.8 per cent. Furthermore, the early starters are highly concentrated in September, with the overall proportion being up to 16 per cent in that particular month, and falling to only 3 per cent from October to December, with a further decline to less than 1 per cent in later months. In contrast, late starters are evenly distributed across all months. Thus, late starters appear to pose a negligible threat to our estimations, which are mainly reliant on a comparison between children born in August and those born in September and other months.²⁵ We therefore go on to evaluate the impact of the inclusion of early starters in our estimates of the MOB effects, hypothesizing a two-fold effect from their inclusion.

Firstly, since early starters will be the youngest students in the class, they are therefore subject to adverse impacts in educational achievement, either through the relative age effect or the absolute age effect. As early starters are clustered in September, this disadvantage will cause a downward bias in our RDD estimate. Secondly, early starters may

²⁵ The almost uniform distribution of late starters across all months implies that redshirting is rare, since the occurrences of redshirting, if present, should be concentrated on those children born in August.

have a second chance to take the CET by the time they reach the age of 18, thereby enhancing their chances of being admitted into a university at the age of 18. This causes an upward bias in our major estimates.

We consider the existence of early starters as a minor concern, for the following reasons. Firstly, since the two impacts of early starters are in opposite directions, they may well cancel out each other. Secondly, the proportion of early starters is fairly low for November and the later months of the school year; however, both our FE and OLS estimates indicate a significant MOB effect for November, with the effect declining monotonically in later months. By extrapolating this pattern, we can reasonably infer that the MOB effects for September and October should also be significant and that they likely be larger than those for November. Indeed, our empirical results confirm this extrapolation. Finally, and perhaps most importantly, the potential bias caused by early starters applies to both our FE and OLS estimates. Under the assumption that the two estimates are biased uniformly in terms of magnitude, the FE-OLS disparity remains unchanged; thus, the main conclusions of our study remain intact.

Sibling samples versus non-sibling samples

We assess the representativeness of our sibling samples by comparing our OLS estimates to those obtained from the non-sibling samples. The results are summarized in Table 6, with columns (1) and (2) reporting the respective estimates of the MOB effects on the male sibling and non-sibling samples.

The p -values from the SURE tests shown in column (3) of Table 6 reveal an insignificant difference between the sibling and non-sibling estimates for the same month, with the only exception of February. The comparisons of the female samples, in the right-hand panel, also point to the same conclusion. Whilst these test results confirm that the estimates based on the two different samples are comparable, the similarity only holds with regard to the magnitude of the estimates, and not with regard to the percent change.

The figures in columns (1) and (2) in the top row of Table 6 show that in the non-sibling sample, the probability of university attendance for males born in August, at 13.0 per cent, is higher than the probability in the sibling sample for males born in August, at 10.7

per cent. Therefore, in terms of the percentage change, the MOB effects are generally smaller for the non-sibling samples than for the sibling samples.

Simulation test

In this section, we design a simulation which enables us to investigate whether parents have preferences for certain months or seasons for the timing of their births, which could lead to the centralization of dates of birth for all siblings.

In this simulation, we select individuals with at least one sibling (sister or brother). We calibrate the gap between the dates of birth of each pair of siblings, which is measured by the absolute value of the calendar gap (in days) between the two dates of birth, regardless of their birth years. If the gap, denoted by d , is higher than 182 days, we replace it with $365-d$. Based upon this measure, we find that the average gap between the dates of birth of the siblings is approximately 86.8 days (S.D. = 53.5 days). These dates of birth are then scrambled and randomly distributed (without replacement) back to all individuals, without altering mother, child or sibling relationships. From this simulated distribution of dates of birth, although the seasonal pattern of births remains, any parental preference for birth timings is eliminated. If parents really do have timing preferences for the birth season or birth month of their children, the average gap between the dates of birth of the siblings based on the simulated data should be larger than the gap based on their actual dates of birth.

Our simulated average gap is around 91.2 days (S.D. = 52.6 days), which is somewhat larger than the actual average gap, although the two figures do not significantly differ according to the results of a formal t test on the equivalence of the two means. Thus, our simulation results suggest that there is no indication of parental selection of the birth month or birth season of children based on parental preferences.

6. Conclusions

We examine the potential endogeneity of the month of birth (MOB) using a sample of approximately one million siblings in Taiwan, with both our FE and RDD results suggesting only a limited role of family background characteristics in determining the MOB effects on the likelihood of university attendance at the age of 18. Our findings are at odds with the US

evidence provided by Buckles and Hungerman (2013), who found that controlling for family background characteristics eliminated about half of the correlation between the birth season and educational achievement (and earnings). This therefore raises the question as to why family characteristics seem to matter in the US, but not in Taiwan.

Buckles and Hungerman (2013) argued that the role played by family characteristics was attributable, at least in some part, to high-SES women disproportionately planning births away from the winter months. As a result, births to high- and low-SES mothers exhibit different seasonal patterns, with high-SES mothers having fewer births in the winter and more births in the spring. This is not, however, the case in Taiwan, since our empirical results suggest strongly similar seasonal patterns for high- and low-SES mothers.

We can only speculate that weather may provide an explanation for the different scenarios between the US and Taiwan. Unlike the US, where the winter can be very cold (indeed, particularly harsh in the northern states), the temperature during a Taiwanese winter is generally mild, ranging from 55° to 70° Fahrenheit. Thus, giving birth during the winter may not constitute a serious concern for Taiwanese mothers contemplating birth plans, regardless of their SES backgrounds. Further evidence is, however, required in order to confirm our supposition.

References

- Angrist, Joshua D. and Alan B. Krueger (1991), "Does Compulsory School Attendance Affect Schooling and Earnings?" *Quarterly Journal of Economics*, 106(4): 979-1014.
- _____ (1992), "The Effect of Age at School Entry on Educational Attainment: An Application on Instrumental Variables with Moments from Two Samples," *Journal of the American Statistical Association*, 87: 328-336.
- Arkes, Jeremy (2010), "Using Unemployment Rates as Instruments to Estimate Returns to Schooling," *Southern Economic Journal*, 76(3): 711-722.
- Barua, Rashmi and Kevin Lang (2010), "School Entry, Educational Attainment and Quarter of Birth: A Cautionary Tale of LATE," Working Papers Series WP2010-019, Boston University - Department of Economics.
- Bedard, Kelly and Elizabeth Dhuey (2012), "School-Entry Policies and Skill Accumulation across Directly and Indirectly Affected Individuals," *Journal of Human Resources*, 47(3): 643-683.
- _____ (2006), "The Persistence of Early Childhood Maturity: International Evidence of Long- Run Age Effects," *The Quarterly Journal of Economics*, 121(4): 1437-1472.
- Bound, John and David A. Jaeger (2000), "Do Compulsory School Attendance Laws Alone Explain the Association Between Quarter of Birth and Earnings?" *Worker Well-Being*, 19: 83-108.
- Bound, John, David A. Jaeger, and Regina M. Baker (1995), "Problems with Instrumental Variables Estimation When the Correlation between the Instruments and the Endogenous Explanatory Variable is Weak," *Journal of the American Statistical Association*, 90: 443-450.
- Buckles, Kasey S. and Daniel M. Hungerman (2013), "Season of Birth and Later Outcomes: Old Questions, New Answers," *Review of Economics and Statistics*, 95(3): 711-724.
- Chernozhukov, Victor and Christian Hansen (2006), "Instrumental Quantile Regression Inference for Structural and Treatment Effect Models," *Journal of Econometrics*, 132(2): 491-525.
- Costa, Dora L. and Joanna N. Lahey (2005) "Predicting Older Age Mortality Trends," *Journal of the European Economic Association*, 3(2-3): 287-293.
- Crawford, Claire, Lorraine Dearden and Costas Meghir (2010), "When You Are Born Matters: the Impact of Date of Birth on Educational Outcomes in England," IFS Working Papers W10/06, Institute for Fiscal Studies.
- Dobkin, Carlos and Fernando Ferreira (2010), "Do School Entry Laws Affect Educational Attainment and Labor Market Outcomes?" *Economics of Education Review*, 29(1): 40-54.
- Heckman, James, Junjian Yi and Junsen Zhang (2013), "Early Health Shocks, Intrahousehold Resource Allocation, and Child Human Capital," manuscript.

- Imbens, Guido and Karthik Kalyanaraman (2012), "Optimal Bandwidth Choice for the Regression Discontinuity Estimator," *The Review of Economic Studies*, 79(3): 933-959.
- Lam, David A. and Jeffery A. Miron (1991), "Seasonality of Births in Human Populations," *Biodemography and Social Biology*, 38(1-2): 51-78.
- Lee, Chanyoung, and Peter Orazem (2010), "High School Employment, School Performance, and College Entry," *Economics of Education Review*, 29(1): 29-39.
- Lee, David S. and David Card (2008), "Regression Discontinuity Inference with Specification Error," *Journal of Econometrics*, 142(2): 655-674.
- Lee, David S. and Thomas Lemieux (2010), "Regression Discontinuity Designs in Economics," *Journal of Economic Literature*, 48(2): 281-355.
- Lefgren, Lars and Frank McIntyre (2006), "The Relationship between Women's Education and Marriage Outcomes," *Journal of Labor Economics*, 24(4): 787-830.
- Lemieux, Thomas and Kevin Milligan (2008), "Incentive Effects of Social Assistance: A Regression Discontinuity Approach," *Journal of Econometrics*, 142(2): 807-828.
- Leigh, Andrew and Chris Ryan (2008), "Estimating the Returns to Education Using Different Natural Experiment Techniques," *Economics of Education Review*, 27(2): 149-160.
- Lokshin, Michael and Sergiy Radyakin (2012), "Month of Birth and Children's Health in India," *Journal of Human Resources*, 47(1): 174-203.
- Maurin, Eric and Julie Moschion (2009), "The Social Multiplier and Labor Market Participation of Mothers," *American Economic Journal: Applied Economics*, 1(1), 251-272.
- McCrary, Justin (2008), "Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test," *Journal of Econometrics*, 142(2): 698-714.
- Puhani, Patrick A. and Andrea M. Weber (2007), "Does the Early Bird Catch the Worm? Instrumental Variable Estimates of Early Educational Effects of Age of School Entry in Germany," *Empirical Economics*, 32: 359-386.
- Robertson, Erin (2011), "The Effects of Quarter of Birth on Academic Outcomes at the Elementary School Level," *Economics of Education Review*, 30(2): 300-311.
- van Hanswijck de Jonge, Laurence, Glenn Waller, and Nicolas Stettler (2003), "Ethnicity Modifies Seasonal Variation in Birth Weight and Weight Gain of Infants," *Journal of Nutrition*, 133(5): 1415-1418.
- Weber, Gerhard W., Hermann Prossinger, and Horst Seidler (1998), "Height Depends on Month of Birth," *Nature*, 391: 754-755.

Table 1: Summary statistics

	Males		Females	
	Mean	Standard deviation	Mean	Standard deviation
% univ. attending	0.13	0.33	0.14	0.35
% pub. univ. attending	0.05	0.21	0.05	0.22
Brith weight (kg)	3.29	0.48	3.20	0.46
Mother's age at birth	24.66	3.49	24.72	3.54
Father's age at birth	27.95	4.29	28.04	4.43
Number of siblings	2.84	0.81	3.36	0.95
Father's education				
College and above	0.05	0.21	0.04	0.20
Vocational college	0.02	0.13	0.02	0.13
High school	0.07	0.26	0.07	0.26
Vocational high school	0.21	0.41	0.21	0.41
Junior high school	0.22	0.42	0.22	0.42
Primary school and below	0.42	0.49	0.43	0.49
Unknown	0.01	0.08	0.01	0.09
Mother's education				
College and above	0.02	0.14	0.02	0.14
Vocational college	0.01	0.10	0.01	0.10
High school	0.04	0.21	0.04	0.21
Vocational high school	0.18	0.38	0.17	0.38
Junior high school	0.24	0.43	0.24	0.43
Primary school and below	0.50	0.50	0.51	0.50
Unknown	0.00	0.03	0.00	0.03
School year cohort				
1978.9-1979.8	0.14	0.34	0.13	0.34
1979.9-1980.8	0.16	0.37	0.16	0.37
1980.9-1981.8	0.17	0.38	0.17	0.38
1981.9-1982.8	0.17	0.37	0.17	0.37
1982.9-1983.8	0.14	0.34	0.14	0.35
1983.9-1984.8	0.12	0.32	0.13	0.33
Observatons	451,755		424,133	

Table 2A: OLS and fixed-effects results for male siblings

	Any university attendance			Public university attendance		
mean(Y August)	0.107			0.042		
	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	FE	p-value for testing on (1)=(2)	OLS	FE	p-value for testing on (1)=(2)
(benchmark: August)						
September	0.0344*** (0.0023)	0.0368*** (0.0028)	0.2942	0.0091*** (0.0015)	0.0107*** (0.0019)	0.2794
October	0.0309*** (0.0023)	0.0307*** (0.0027)	0.9554	0.0086*** (0.0015)	0.0085*** (0.0019)	0.9097
November	0.0303*** (0.0023)	0.0302*** (0.0028)	0.9464	0.0098*** (0.0015)	0.0099*** (0.0019)	0.9485
December	0.0260*** (0.0023)	0.0277*** (0.0028)	0.4818	0.0086*** (0.0015)	0.0104*** (0.0019)	0.2429
January	0.0210*** (0.0023)	0.0224*** (0.0028)	0.5540	0.0077*** (0.0015)	0.0092*** (0.0019)	0.3200
February	0.0165*** (0.0024)	0.0162*** (0.0029)	0.8996	0.0060*** (0.0015)	0.0056*** (0.0020)	0.8068
March	0.0188*** (0.0024)	0.0163*** (0.0029)	0.2968	0.0067*** (0.0015)	0.0059*** (0.0020)	0.6287
April	0.0155*** (0.0024)	0.0123*** (0.0029)	0.1988	0.0058*** (0.0015)	0.0044** (0.0020)	0.3802
May	0.0040* (0.0023)	0.0062** (0.0029)	0.3648	0.0018 (0.0015)	0.0026 (0.0020)	0.5788
June	0.0073*** (0.0023)	0.0067** (0.0029)	0.8248	0.0027* (0.0015)	0.0020 (0.0019)	0.6939
July	0.0012 (0.0022)	-0.0003 (0.0028)	0.5249	0.0015 (0.0014)	0.0017 (0.0019)	0.8842
All MOB coefficients			0.4171			0.6171
Constant	0.0364*** (0.0039)	0.0371*** (0.0056)		451755 0.0008	451755 0.0008	
Birth year cohorts	yes	yes		yes	yes	
Birth weight	yes	yes		yes	yes	
Number of individuals: 451,755						
Number of family: 216,765						

Note: (1) Observations are children born during 1978/01/01 and 1984/12/31. (2) The male (female) sample are children who have at least one male (female) sibling; (3) Parental education controls include a full set of indicators for education levels; (4) Robust standard errors in parentheses for OLS estimations, which are clustered at the family level; (5) *** p<0.01, ** p<0.05, * p<0.1.

Table 2B: OLS and fixed-effects results for female siblings

	Any university attendance			Public university attendance		
mean(Y August)	0.119			0.046		
	(1)	(2)	(3)	(4)	(5)	(3)
	OLS	FE	p-value for test on (1)=(2)	OLS	FE	p-value for test on (4)=(5)
(benchmark: August)						
September	0.0285*** (0.0025)	0.0301*** (0.0029)	0.5115	0.0063*** (0.0016)	0.0053*** (0.0020)	0.5058
October	0.0349*** (0.0024)	0.0313*** (0.0029)	0.1310	0.0086*** (0.0016)	0.0060*** (0.0019)	0.0883
November	0.0369*** (0.0025)	0.0317*** (0.0029)	0.0325	0.0120*** (0.0016)	0.0089*** (0.0020)	0.0463
December	0.0319*** (0.0025)	0.0301*** (0.0030)	0.4734	0.0070*** (0.0016)	0.0057*** (0.0020)	0.3826
January	0.0263*** (0.0025)	0.0213*** (0.0030)	0.0467	0.0094*** (0.0016)	0.0058*** (0.0020)	0.0205
February	0.0233*** (0.0026)	0.0202*** (0.0031)	0.2397	0.0072*** (0.0017)	0.0063*** (0.0021)	0.5734
March	0.0224*** (0.0026)	0.0182*** (0.0030)	0.1014	0.0072*** (0.0016)	0.0023 (0.0020)	0.0028
April	0.0182*** (0.0026)	0.0134*** (0.0031)	0.0645	0.0059*** (0.0017)	0.0020 (0.0021)	0.0173
May	0.0164*** (0.0025)	0.0078** (0.0031)	0.0008	0.0054*** (0.0016)	0.0019 (0.0020)	0.0321
June	0.0096*** (0.0025)	0.0067** (0.0030)	0.2479	0.0024 (0.0016)	0.0004 (0.0020)	0.2032
July	0.0033 (0.0024)	0.0027 (0.0029)	0.7824	-0.0002 (0.0015)	-0.0021 (0.0020)	0.2212
All MOB coefficients			0.0104			0.1099
Constant	0.0416*** (0.0043)	0.0417*** (0.0059)		0.0155*** (0.0027)	0.0143*** (0.0040)	
Birth year cohorts	yes	yes		yes	yes	
Birth weight	yes	yes		yes	yes	
Number of individuals: 424,133						
Number of family: 196,033						

Note: (1) Observations are children born during 1978/01/01 and 1984/12/31. (2) The male (female) sample are children who have at least one male (female) sibling; (3) Parental education controls include a full set of indicators for education levels; (4) Robust standard errors in parentheses for OLS estimations, which are clustered at the family level; (5) *** p<0.01, ** p<0.05, * p<0.1.

Table 3: Estimated discontinuities in observable characteristics

	Father's yrs of schooling		Mother's yrs of schooling		Father's age		Mother's age	
	Males	Females	Males	Females	Males	Females	Males	Females
T (=1 if born on or after Sept. 1st)	0.0215 (0.0376)	0.0553 (0.0484)	-0.0382 (0.0425)	0.0472 (0.0402)	-0.0088 (0.0663)	-0.0176 (0.0768)	-0.0483 (0.0550)	-0.0584 (0.0511)
F (standardized birthdate)	0.0014 (0.0012)	-0.0024* (0.0014)	0.0022 (0.0016)	-0.0017 (0.0011)	-0.0041 (0.0026)	-0.0007 (0.0030)	0.0003 (0.0018)	0.0009 (0.0023)
T*F	-0.0004 (0.0017)	0.0028 (0.0019)	0.0012 (0.0021)	0.0032** (0.0015)	0.0036 (0.0040)	-0.0003 (0.0040)	-0.0005 (0.0029)	0.0005 (0.0028)
Constant	9.0492*** (0.0238)	8.9616*** (0.0349)	8.1497*** (0.0280)	8.0676*** (0.0314)	27.7062*** (0.0488)	27.8184*** (0.0568)	24.4481*** (0.0355)	24.4957*** (0.0401)
Optimal bandwidth	40.94	41.54	38.3	44.03	30.14	31.32	30.05	30.87
Observations	104,514	99,237	99,597	106,653	79,068	75,935	79,068	73,491
	Percent first-born child		Birth weigh (kg)		No. of siblings		Percent urban	
	Males	Females	Males	Females	Males	Females	Males	Females
T (=1 if born on or after Sept. 1st)	-0.0023 (0.0065)	0.0088** (0.0039)	-0.0012 (0.0058)	-0.0033 (0.0049)	0.0079 (0.0081)	-0.0060 (0.0119)	-0.0048 (0.0049)	0.0072 (0.0059)
F (standardized birthdate)	0.0000 (0.0001)	0.0001** (0.0000)	0.0001 (0.0002)	-0.0002*** (0.0001)	0.0002 (0.0003)	-0.0002 (0.0002)	0.0000 (0.0001)	-0.0002* (0.0001)
T*F	0.0003 (0.0002)	-0.0003*** (0.0001)	-0.0000 (0.0002)	0.0006*** (0.0001)	-0.0008** (0.0004)	0.0001 (0.0003)	0.0001 (0.0002)	0.0003 (0.0002)
Constant	0.4027*** (0.0040)	0.3957*** (0.0029)	3.2692*** (0.0046)	3.1725*** (0.0036)	2.8381*** (0.0051)	3.3580*** (0.0078)	0.3139*** (0.0035)	0.3055*** (0.0046)
Optimal bandwidth	48.11	109.62	48.19	80.95	38.82	62.7	54.88	56.41
Observations	125,431	254,176	125,431	192,002	99,647	150,169	140,734	135,683

Note: (1) Optimal bandwidths are determined using the procedure proposed by Imbens and Kalyanaraman (2012); (2) Robust standard errors in parentheses, which are clustered at the day level; (3) *** p<0.01, ** p<0.05, * p<0.1.

Table 4: Results from regression discontinuity design

Dependant variable:	Any university attendance				Public university attendance			
	Group level	Individual level			Group level	Individual level		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: male siblings								
mean(Y August)		0.101				0.040		
Optimal bandwidth		39.2				45.2		
T (=1 if born on or after Sept. 1st)	0.0329*** (0.0036)	0.0329*** (0.0035)	0.0330*** (0.0036)	0.0351*** (0.0031)	0.0094*** (0.0019)	0.0094*** (0.0019)	0.0094*** (0.0019)	0.0101*** (0.0018)
F (standardized birthdate)	0.0001 (0.0001)	0.0001 (0.0001)	0.0002 (0.0001)	0.0001 (0.0001)	0.0000 (0.0000)	0.0000 (0.0000)	0.0000 (0.0000)	0.0000 (0.0000)
T*F	0.0002 (0.0002)	0.0002 (0.0002)	0.0001 (0.0002)	0.0001 (0.0002)	0.0001 (0.0001)	0.0001 (0.0001)	0.0001 (0.0001)	0.0000 (0.0001)
County of birth and birth cohorts	--	No	Yes	Yes	--	No	Yes	Yes
Family and ind. characteristics	--	No	No	Yes	--	No	No	Yes
Observations	79	102,330	102,330	102,330	91	117,724	117,724	117,724
Panel B: female siblings								
mean(Y August)		0.111				0.045		
Optimal bandwidth		37.7				45.3		
T (=1 if born on or after Sept. 1st)	0.0354*** (0.0054)	0.0354*** (0.0053)	0.0356*** (0.0051)	0.0351*** (0.0046)	0.0060** (0.0025)	0.0060** (0.0024)	0.0060** (0.0024)	0.0057** (0.0023)
F (standardized birthdate)	-0.0001 (0.0002)	-0.0001 (0.0002)	-0.0001 (0.0002)	-0.0001 (0.0002)	-0.0000 (0.0001)	-0.0000 (0.0001)	-0.0000 (0.0001)	0.0000 (0.0001)
T*F	0.0003 (0.0002)	0.0003 (0.0002)	0.0003 (0.0002)	0.0002 (0.0002)	0.0001 (0.0001)	0.0001 (0.0001)	0.0001 (0.0001)	0.0000 (0.0001)
County of birth and birth cohorts	--	Yes	No	Yes	--	No	Yes	Yes
Family and ind. characteristics	--	No	No	Yes	--	No	No	Yes
Observations	75	90,091	90,091	90,091	91	109,089	109,089	109,089

Note: (1) Optimal bandwidths are determined using the procedure proposed by Imbens and Kalyanaraman (2012); (2) Robust standard errors in parentheses, which are clustered at the day level; (3) *** p<0.01, ** p<0.05, * p<0.1.

Table 5: Exploring the seasonality of mother's education

	(1) females	(2) males
mean(Y August)	0.246	0.252
(benchmark: August)		
September	0.0045 (0.0031)	0.0027 (0.0030)
October	-0.0010 (0.0032)	0.0020 (0.0031)
November	0.0023 (0.0032)	0.0085*** (0.0031)
December	0.0013 (0.0032)	0.0079** (0.0032)
January	0.0018 (0.0032)	0.0029 (0.0031)
February	0.0001 (0.0031)	0.0013 (0.0031)
March	-0.0029 (0.0030)	-0.0008 (0.0030)
April	-0.0046 (0.0030)	-0.0023 (0.0030)
May	0.0009 (0.0030)	0.0019 (0.0029)
June	0.0024 (0.0030)	0.0038 (0.0030)
July	0.0000 (0.0031)	0.0079*** (0.0030)
Constant	0.7142*** (0.0073)	0.7366*** (0.0070)
Birth year cohorts	yes	yes
Birth weight	yes	yes
county of residence	yes	yes
sibling size	yes	yes
Observations	424,133	451,755

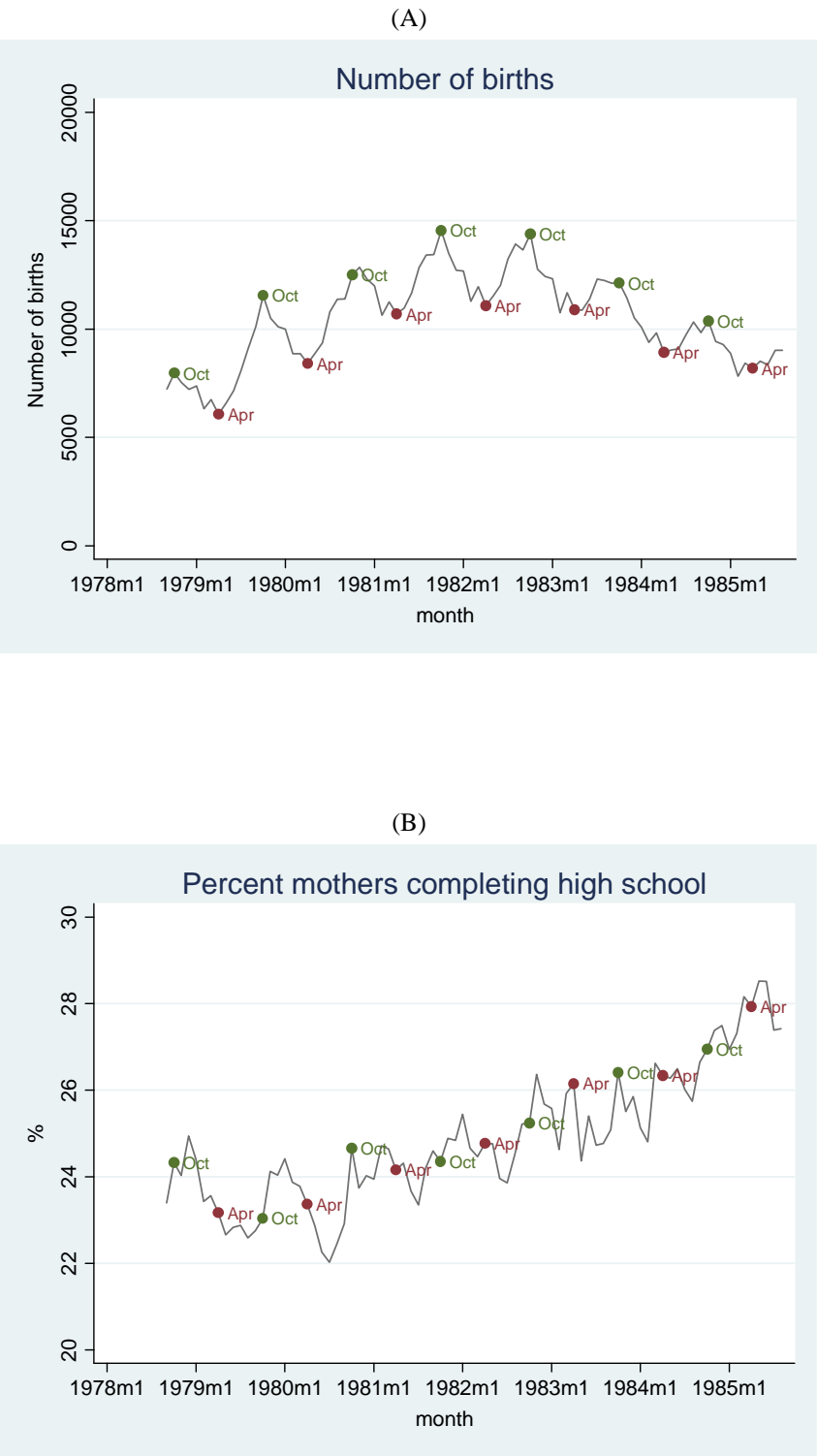
Note: (1) Observations are children born during 1978/01/01 and 1984/12/31. (2) The male (female) sample are children who have at least one male (female) sibling; (3) Robust standard errors in parentheses, which are clustered at the family level. (5)*** p<0.01, ** p<0.05, * p<0.1.

**Table 6: Comparing the MOB effects on university attendance
using sibling and non-sibling samples**

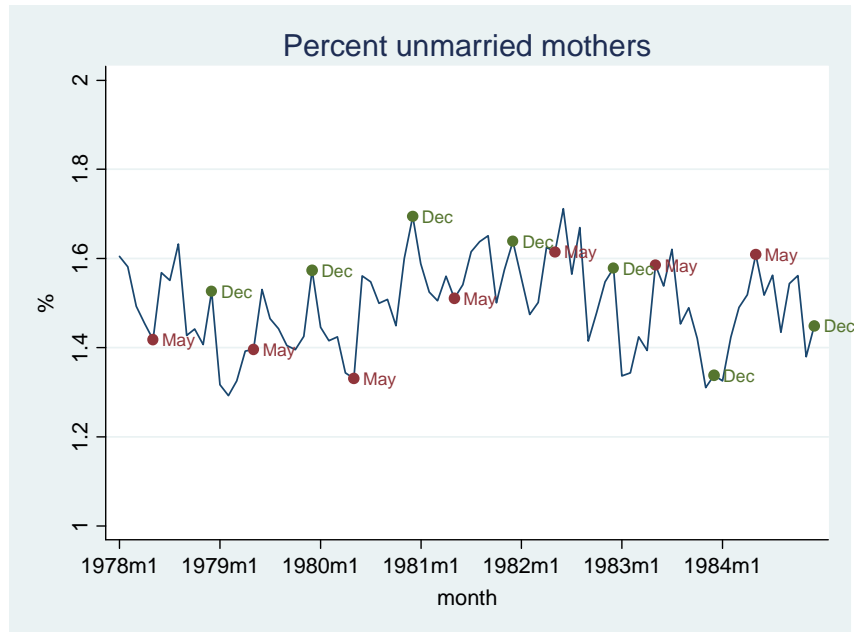
mean(Y August)	males			females		
	0.107	0.138		0.119	0.161	
	(1)	(2)	(3)	(4)	(5)	(6)
	Sibling sample	Non-sibling sample	p-value for testing on (1)=(2)	Sibling sample	Non-sibling sample	p-value for testing on (1)=(2)
(benchmark: August)						
September	0.0344*** (0.0023)	0.0350*** (0.0023)	0.8633	0.0285*** (0.0025)	0.0297*** (0.0025)	0.7274
October	0.0309*** (0.0023)	0.0360*** (0.0023)	0.1098	0.0349*** (0.0024)	0.0362*** (0.0025)	0.7108
November	0.0303*** (0.0023)	0.0329*** (0.0023)	0.4402	0.0369*** (0.0025)	0.0338*** (0.0025)	0.3863
December	0.0260*** (0.0023)	0.0288*** (0.0023)	0.3922	0.0319*** (0.0025)	0.0346*** (0.0026)	0.4564
January	0.0210*** (0.0023)	0.0244*** (0.0023)	0.2967	0.0263*** (0.0025)	0.0276*** (0.0026)	0.7195
February	0.0165*** (0.0024)	0.0237*** (0.0024)	0.0352	0.0233*** (0.0026)	0.0277*** (0.0026)	0.2359
March	0.0188*** (0.0024)	0.0191*** (0.0023)	0.9405	0.0224*** (0.0026)	0.0214*** (0.0026)	0.7774
April	0.0155*** (0.0024)	0.0168*** (0.0024)	0.6982	0.0182*** (0.0026)	0.0205*** (0.0026)	0.5255
May	0.0040* (0.0023)	0.0084*** (0.0023)	0.1761	0.0164*** (0.0025)	0.0130*** (0.0026)	0.3512
June	0.0073*** (0.0023)	0.0046** (0.0023)	0.4072	0.0096*** (0.0025)	0.0046* (0.0025)	0.1564
July	0.0012 (0.0022)	0.0053** (0.0022)	0.1966	0.0033 (0.0024)	0.0005 (0.0024)	0.4087
All MOB coefficients			0.2798			0.3232
Constant	0.0364*** (0.0039)	0.0386*** (0.0040)		0.0416*** (0.0043)	0.0604*** (0.0044)	
Birth year cohorts	yes	yes		yes	yes	
Birth weight	yes	yes		yes	yes	
Observations	451,755	535,583		424,133	499,582	

Note: (1) Observations are children born during 1978/01/01 and 1984/12/31. (2) The male (female) sample are children who have at least one male (female) sibling; (3) Parental education controls include a full set of indicators for education levels; (4) Robust standard errors in parentheses for OLS estimations, which are clustered at the family level; (5)*** p<0.01, ** p<0.05, * p<0.1.

Figure 1: Seasonality of number of births and maternal characteristics



(C)



(D)

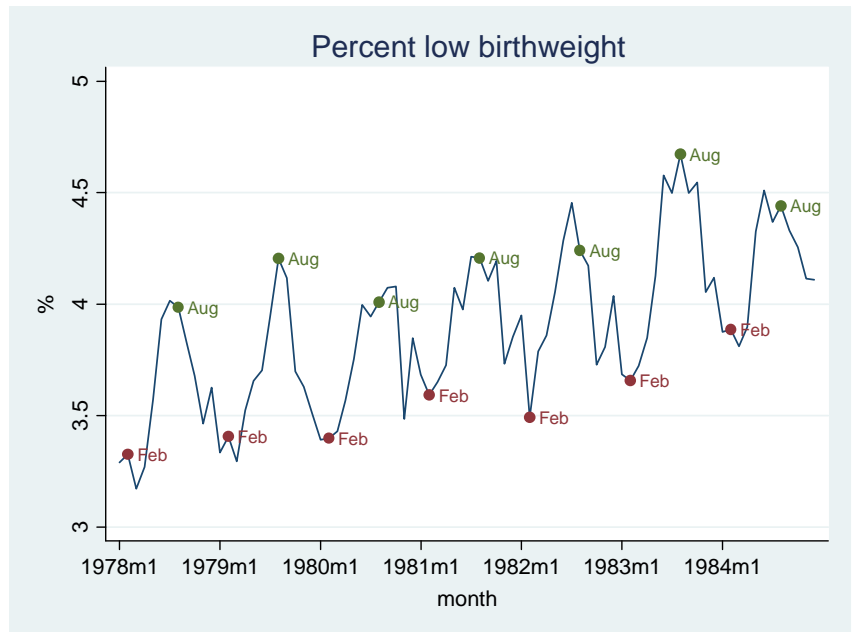
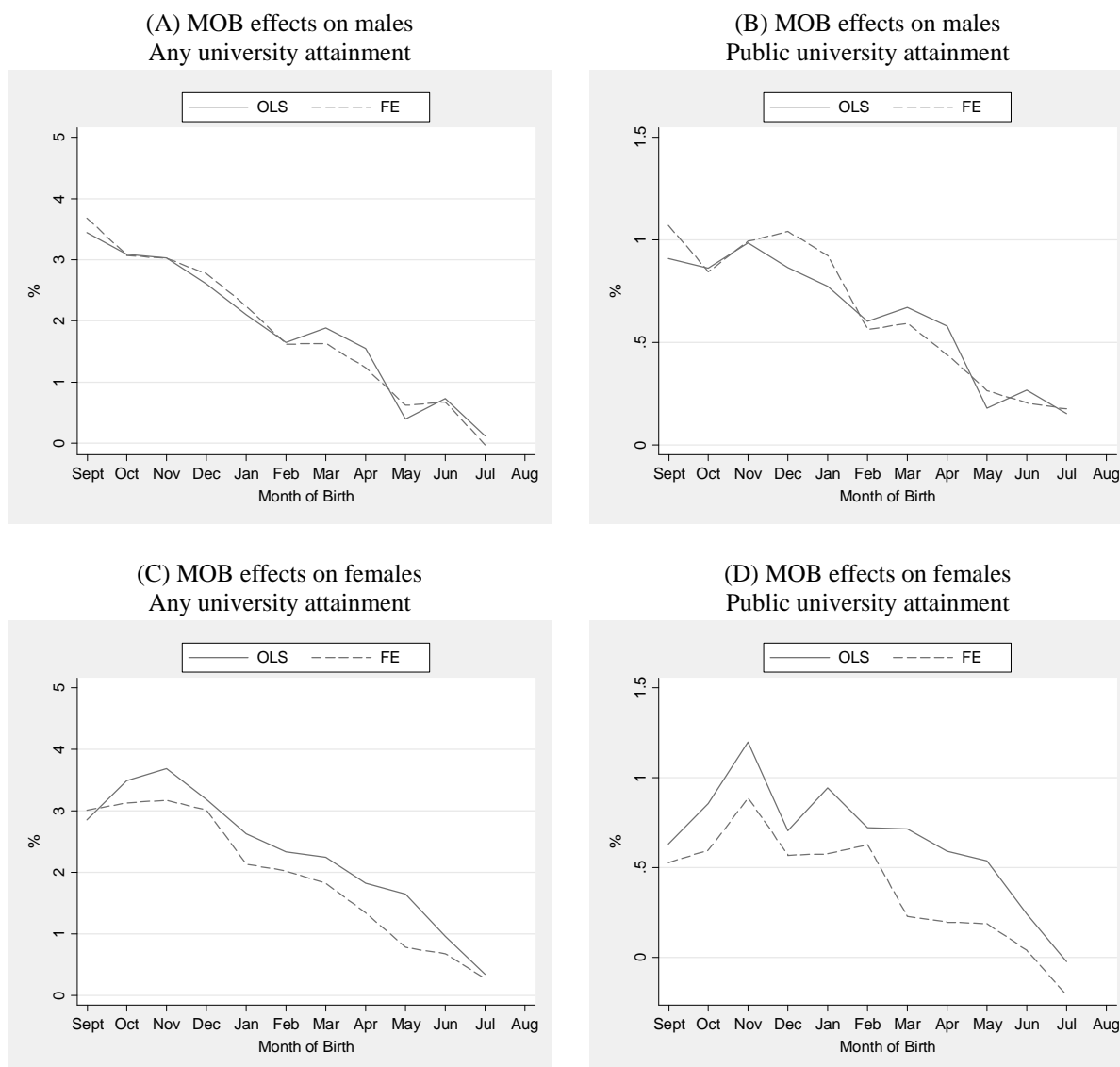


Figure 2: Comparing the FE and OLS estimates



Notes: OLS estimates are obtained from regressions without controlling for family background variables.

Figure 3A: Means of the observables, males

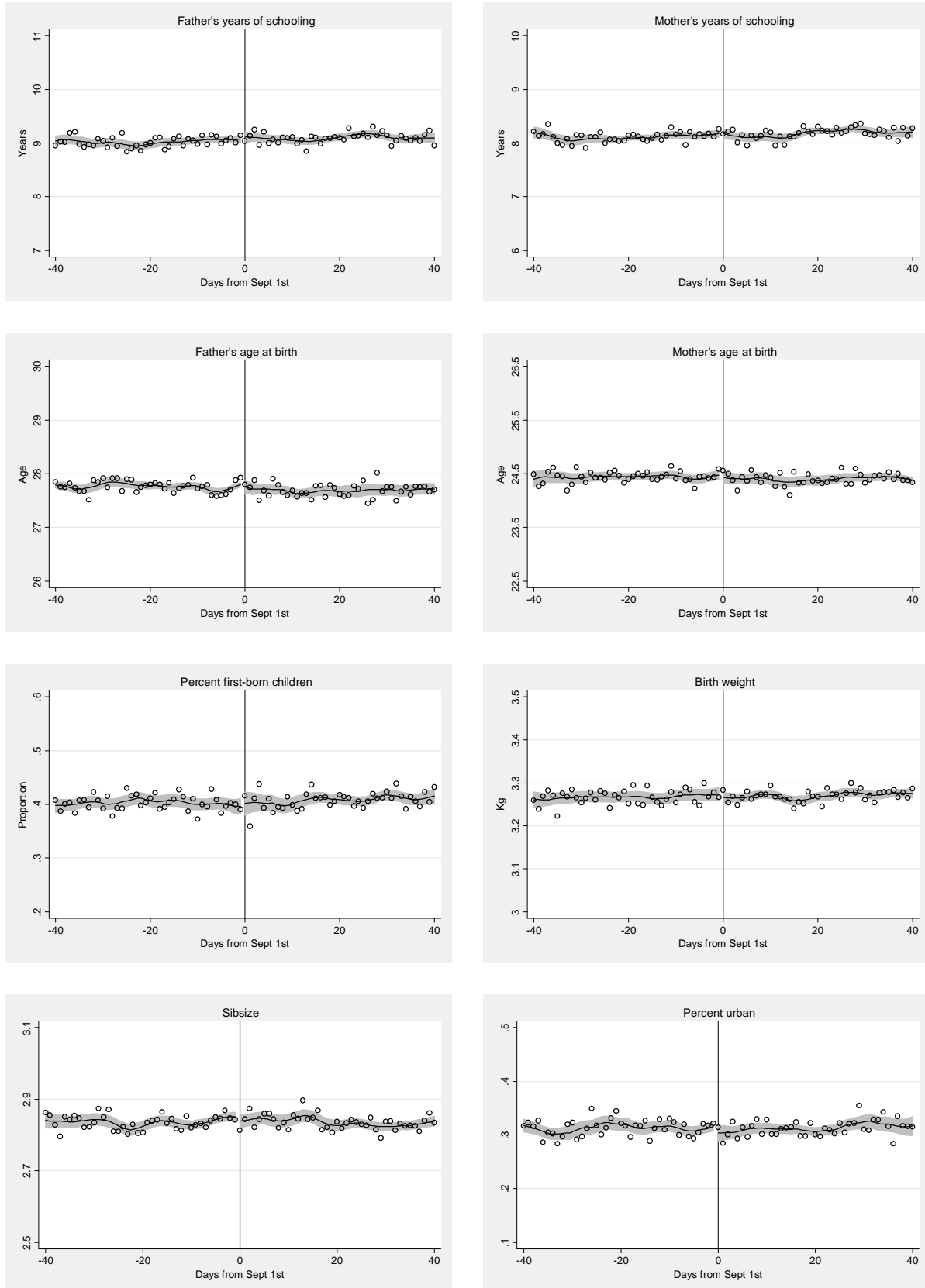


Figure 3B: Means of the observables, females

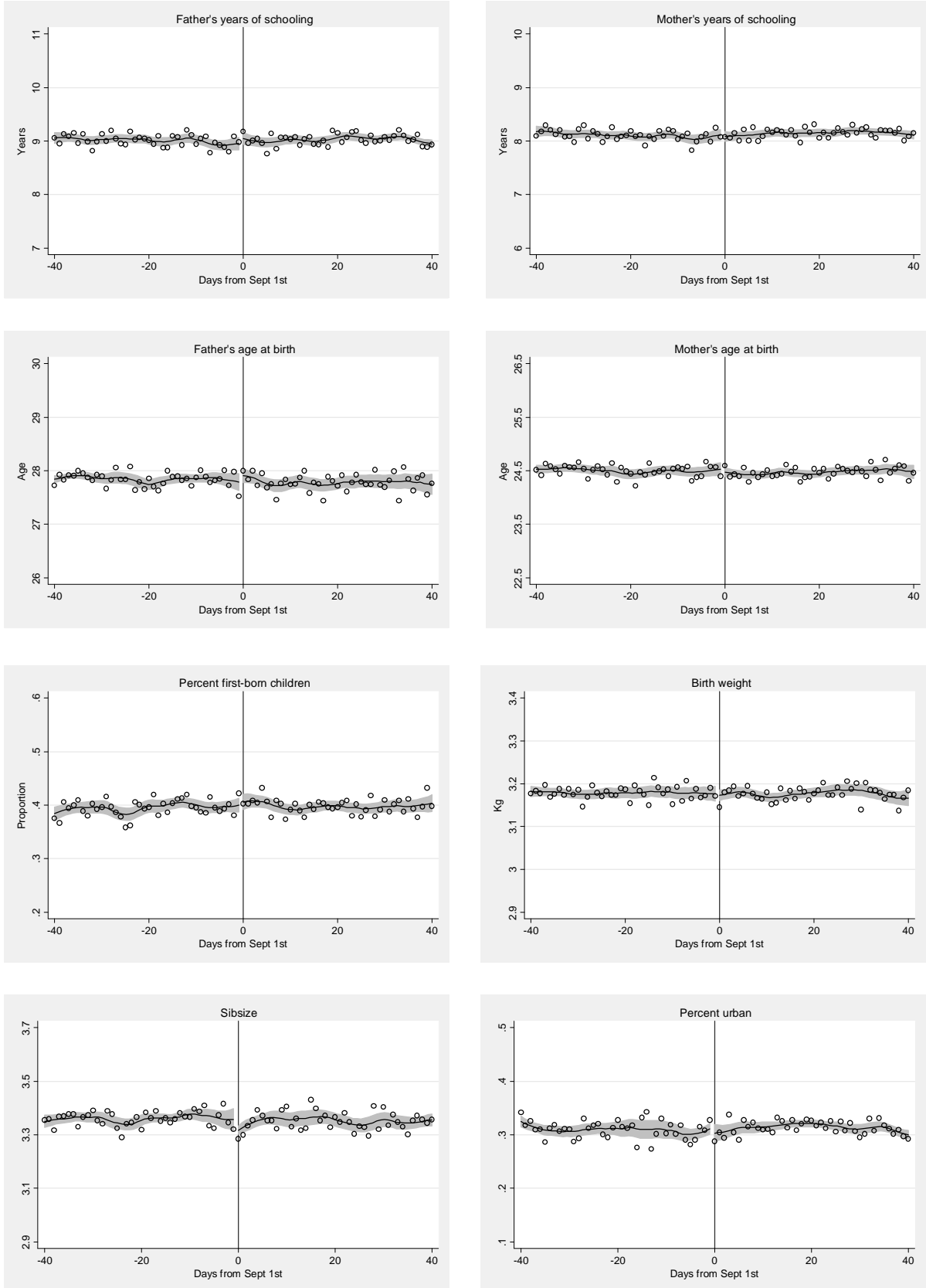
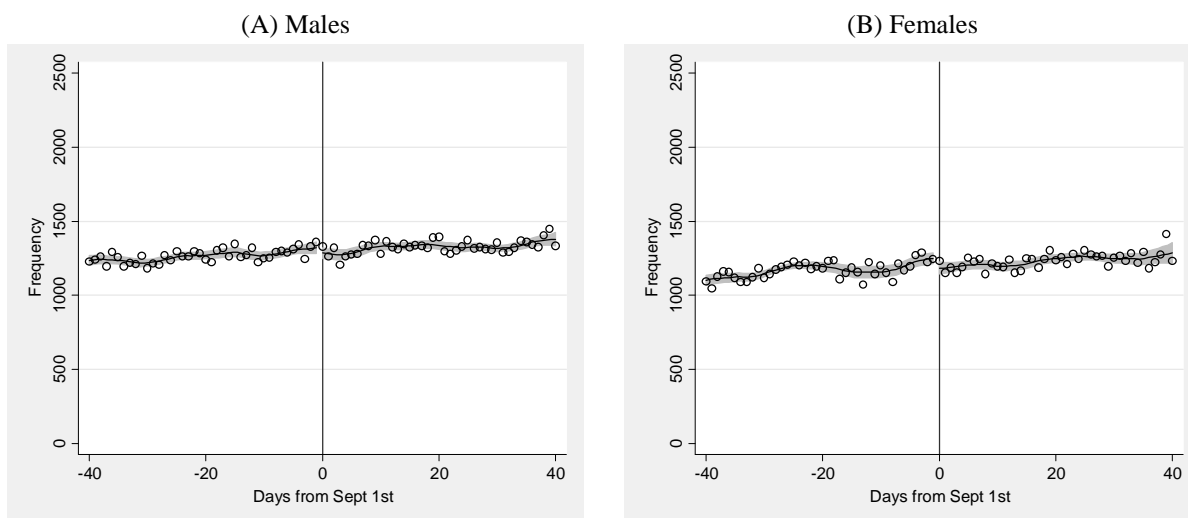
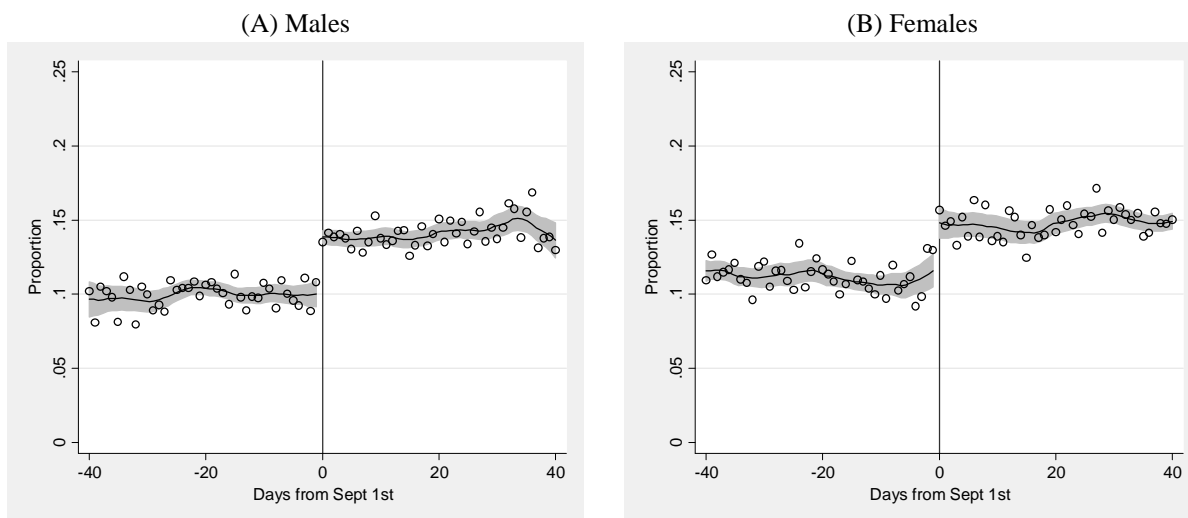


Figure 4: Daily number of births across standardized dates of birth



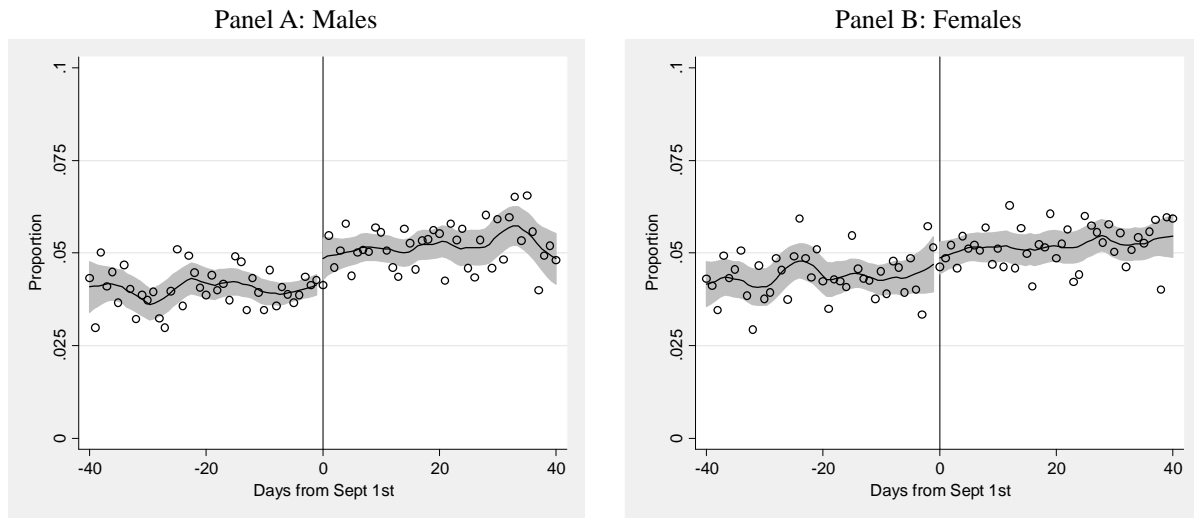
Notes: The X-axis measures the distance in terms of days from the first of September. Grey areas are confidence belts.

Figure 5: Probability of general university attendance for individuals born around the first of September



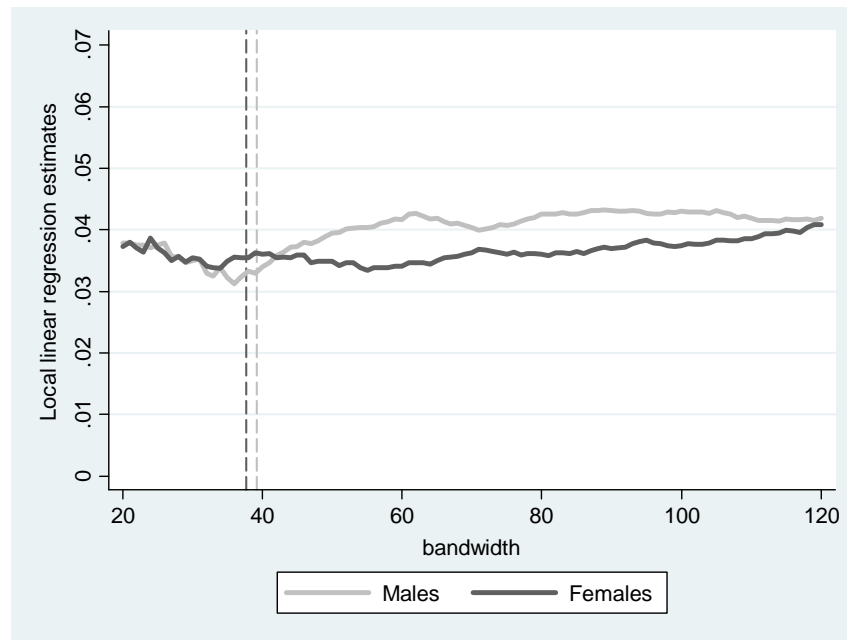
Notes: The X-axis measures the distance in terms of days from the first of September. Grey areas are confidence belts.

Figure 6: Probability of public university attendance for individuals born around the first of September



Notes: The X-axis measures the distance in terms of days from the first of September. Grey areas are confidence belts.

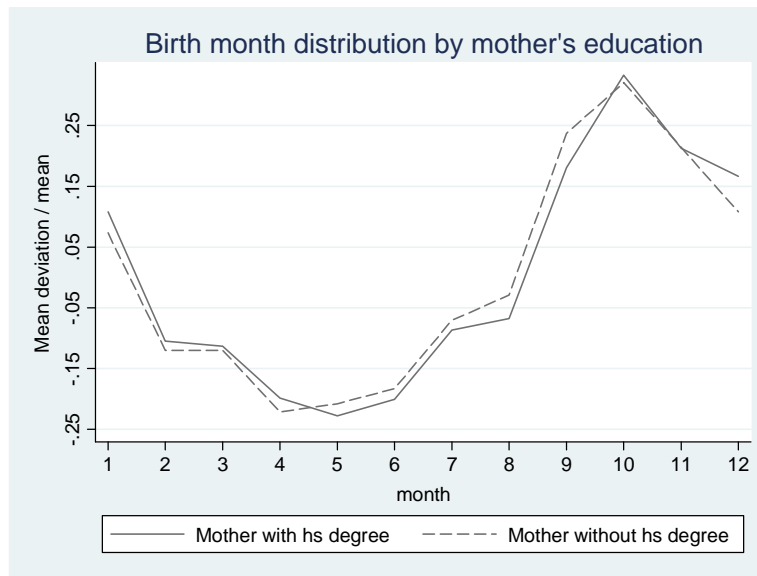
Figure 7: Local linear regression results using different bandwidths



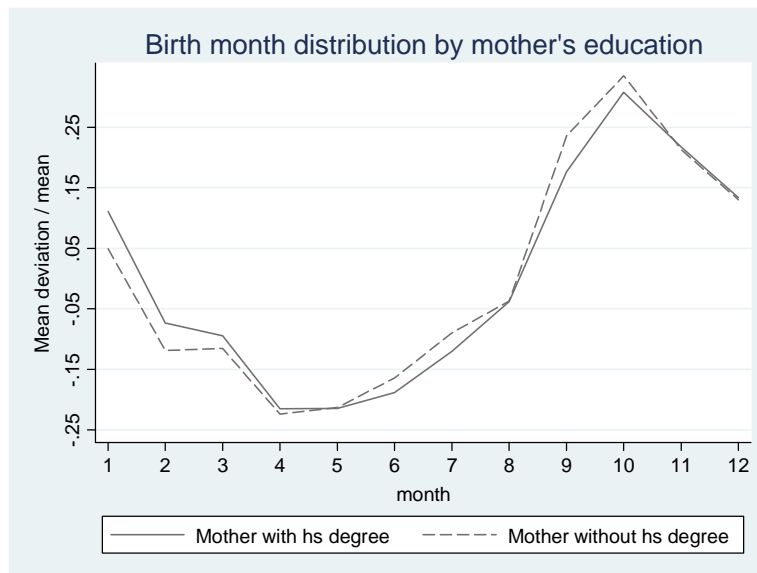
Notes: Estimates are carried out by estimating local linear regressions without controlling for covariates. The darker curve and the lighter curve represent the optimal bandwidths for the male (39.2) and female (37.7) samples, respectively, and both are determined using the procedure proposed by Imbens and Kalyanaraman (2012).

Figure 8: Seasonality of firstborn births for mothers with and without high school education

(A) males



(B) females



Each point in the curves represents a standardized value, which is obtained by having the total number of births in the month subtracted by the average monthly births of the relevant year, and divided by the average.

Appendix

Exploring the role of intra-family resource allocation

One potential threat to our FE model arises from intra-family resources reallocation: parents, as discussed in Section 3.1, may allocate investment across children to reinforce or compensate the gap among children's endowments. The resulting FE estimates of the MOB effects are therefore potentially biased, so our FE-OLS comparisons would fail to reflect the true influences from family background. To evaluate this potential bias, we examine whether the incident of university attendance of the firstborn child is affected by the quarter of birth (QOB) for the second child, given the first child's QOB. Specifically, we select families with two children, despite the genders, and categorize each sibling into the following four quarters of birth based on the school year (not calendar year) as:

- (1) Quarter 1 (denoted by QOB_1) from September to November;
- (2) Quarter 2 (denoted by QOB_2) from December to February;
- (3) Quarter 3 (denoted by QOB_3) from March to May;
- (4) Quarter 4 (denoted by QOB_4) from June to August.

and estimate the following regression for the firstborn children born in each of the four quarters of birth in turn:

$$Uni_i^1 = \gamma_0 + \sum_{k=2}^4 (\gamma_k \cdot QOB_k^2) + \pi_0 \cdot sex^2 + \sum_{k=2}^4 \pi_k (sex^2 \cdot QOB_k^2) + X_i \lambda + \varepsilon_i \quad (6)$$

Where Uni_i^1 is a dummy indicating university attendance at age 18 for the first-born child ($Uni_i^1=1$); QOB_k^2 is a dummy variable indicating the second child being born in quarter k ($k = 2, 3, \text{ or } 4$), leaving the first quarter as the omitted group; sex^2 refers to the gender of the second child ($sex^2=1$ for males; $=0$ for females); and X_i are controls. In equation (6), γ_k captures the effect of the second child being born in quarter k relative to a second child being born in the first (omitted) quarter. With the interaction terms of the second child's QOB and gender, the specification of equation (6) allows γ_k to differ for the second child being a boy or a girl, and π_k captures this additional effect of γ_k for boys relative to girls. If intra-family resource allocation is independent of children's QOBs, all the estimates of γ_k and π_k are expected to be zero.

The estimates from estimating equation (6) are presented in Table A1. Column (1) shows the coefficient estimates for γ_k and π_k for the firstborn females who are born in Q1. The fact that none of the coefficients is significant suggests that neither the QOB nor the

gender of the second child matters for the likelihood of the older sister attending university at age 18. This conclusion holds for firstborn females born in all other three quarters, as shown by columns (2) to (4). These results apparently suggest that the intra-family resource allocation does not constitute a major concern for the validity of our FE estimates of the MOB effects.

The estimates for the males, however, presents a somewhat different story. While all coefficients but one of γ_k and π_k remains as zero in columns (5), (6), and (8), all the estimates of γ_k are significantly positive and the estimates for π_k are significantly negative in column (7). These imply that, for a firstborn male born in Q3, having a sister born in either Q2, Q3, or Q4 will increase his odd to attend a university, and the effects appear to be smaller if the second child is a boy. These patterns, however, are not consistent with either the reinforcement or compensation hypothesis. Other than this exception, we find very little evidence about an active role of intra-family resource allocation in determining MOB effects between siblings.

Table A1: Examining the effects of second child

	Firstborn males				Firstborn females			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
	Fall Born (Sept to Nov)	Winter Born (Dec to Feb)	Spring Born (Mar to May)	Summer born (Jun to Aug)	Fall Born (Sept to Nov)	Winter Born (Dec to Feb)	Spring Born (Mar to May)	Summer born (Jun to Aug)
Q2	0.0077 (0.0057)	0.0045 (0.0065)	0.0142** (0.0070)	0.0049 (0.0062)	0.0019 (0.0066)	-0.0005 (0.0074)	-0.0076 (0.0083)	0.0042 (0.0072)
Q3	0.0205*** (0.0060)	0.0024 (0.0065)	0.0137** (0.0066)	0.0076 (0.0063)	0.0086 (0.0068)	0.0043 (0.0074)	0.0047 (0.0079)	0.0093 (0.0073)
Q4	0.0050 (0.0057)	0.0007 (0.0065)	0.0194*** (0.0065)	-0.0004 (0.0057)	0.0005 (0.0066)	-0.0010 (0.0074)	0.0095 (0.0078)	-0.0029 (0.0065)
brother	0.0003 (0.0053)	-0.0022 (0.0066)	0.0173** (0.0067)	-0.0010 (0.0057)	-0.0086 (0.0061)	-0.0037 (0.0074)	-0.0036 (0.0079)	0.0032 (0.0067)
Q2*brother	-0.0033 (0.0078)	-0.0082 (0.0089)	-0.0150 (0.0098)	-0.0032 (0.0085)	0.0055 (0.0088)	-0.0043 (0.0099)	-0.0044 (0.0111)	-0.0095 (0.0097)
Q3*brother	-0.0079 (0.0081)	-0.0015 (0.0089)	-0.0267*** (0.0091)	0.0044 (0.0087)	-0.0014 (0.0091)	0.0003 (0.0099)	-0.0126 (0.0105)	-0.0109 (0.0098)
Q4*brother	-0.0002 (0.0078)	-0.0018 (0.0089)	-0.0191** (0.0090)	0.0050 (0.0077)	0.0078 (0.0088)	0.0016 (0.0099)	-0.0126 (0.0103)	0.0002 (0.0087)
Observations	65,613	55,490	45,498	50,757	56,801	48,988	40,133	44,810

Note: (1) Control variables are parental education levels, parental ages, total number of siblings, birth order, and birthweight; (2) Robust standard errors in parentheses for OLS estimations, which are clustered at the family level; (3)*** p<0.01, ** p<0.05, * p<0.1.