



The Effect of Age at First Marriage on Female Fertility: Evidence from Korea's School Entry Policy Using Exact Date of Birth

Namho Kwon¹ · Hosung Sohn²

Received: 17 July 2021 / Accepted: 1 November 2022 / Published online: 10 February 2023
© The Author(s), under exclusive licence to Springer Nature B.V. 2023

Abstract

Delay in age at first marriage has been hypothesized as a significant determinant of worldwide fertility decline in high-income countries. Testing this hypothesis is challenging because age at first marriage is confounded with unobservable variables such as career aspiration. To overcome the endogeneity issue in the age at first marriage variable, this paper exploits exact date of birth that creates an exogenous variation in school starting age as an instrument for female's age at first marriage and uses a regression discontinuity design to estimate its impact on female fertility. The statistical analysis based on the 2009 Nationwide Fertility Level and Family Health and Welfare Survey shows that an increase of one year in age at first marriage reduces the likelihood of any childbirth (extensive margin) by about 8 percentage points (10%) and total childbirths (intensive margin) by 0.1 children (6.3%). While delay in age at first marriage reduces fertility, we argue that policymakers should not implement policies to prevent females from delaying marriage timing as many studies have found that there are gains to delaying marriage. Rather, we argue that policymakers should engage more in identifying policy measures that allow the transition to adulthood and participation in higher-education institutions more compatible with motherhood and childbearing.

Keywords Age at first marriage · Fertility · Natalist policy · Regression discontinuity design

JEL Classification N35 · J12 · J13

✉ Hosung Sohn
sohnhosung@gmail.com

Namho Kwon
namhokwon@gmail.com

¹ Department of Public Administration, Soongsil University, Sangdo-ro 369, Dongjak-gu, Seoul 06978, South Korea

² School of Public Service, College of Social Sciences, Chung-Ang University, Hukseok-ro 84, Dongjak-gu, Seoul 06794, South Korea

Introduction

Fertility rates in many high-income countries are well below the replacement rate of 2.1 children, and such rates are often correlated with factors such as low mortality rate, high education level, urbanization, and an increase in female labor participation. In 2015, for example, the total fertility rate (TFR) of South Korea, Japan, Germany, and Spain was 1.4 or less. While optimal replacement fertility varies by country (Espenshade et al., 2003) and the proposition that a TFR well below the replacement rate poses a serious threat to a country is not unanimously agreed upon by researchers, a recent study that analyzed National Transfer Accounts data for 40 countries showed that a moderately low fertility level is favorable for the broader material standard of living (Lee et al., 2014). Nonetheless, many agree that a very low fertility rate will challenge the operation of many government programs and that it will undermine living standards, especially when the low rate is coupled with aging.

Because maintaining a moderate level of fertility is critical for any nation's sustained economic development, many governments facing low birthrates and fast-aging populations are devoting a high share of government expenditure on natalist policies. Despite such engagement, however, birth rates have not rebounded in most countries. For instance, the TFR of South Korea is the lowest among OECD countries even though the Korean government has spent around 100 billion dollars on natalist policies over the last decade.

The fact that the TFR in countries with very low birthrates did not change does not necessarily imply that natalist policies were ineffective. The rate might have declined further if such policies had not been implemented. Nevertheless, a large number of researchers believe that the natalist policies implemented in many of the countries were unsuccessful. We argue that there are two main reasons for the ineffectiveness of many of the natalist policies. First, governments may not have identified the “drivers” of low fertility. For example, if the drivers of low fertility are related to the expenses of fostering, then government policies should aim to reduce such expenses. If the existing policies did not target such expenses, their success in raising fertility was likely to be limited. Second, existing policies may have failed to target the drivers of low fertility even though governments were well aware of the cause. Suppose the drivers of low fertility is females' incapability to achieve work–life balance. In order for females to have childbirths, it is necessary for government policies to eliminate the factors that hinder work–life balance. Work–life balance is subjective, however, and consequently the factors hindering such a balance may not have been targeted properly by government policies, even though the policies were somehow related to work-life balance.

Identifying the drivers of low fertility and targeting such causes is, of course, easier said than done. Furthermore, if there are many causes, it is also necessary to analyze the impact of a cause relative to the others. From a policy perspective, therefore, it is critical that governments engage in identifying the drivers of low fertility properly and their relative impact precisely if they are to develop effective natalist policies. In this study, we contribute to such efforts by analyzing

the causal impact of age at first marriage on female fertility, hypothesizing that late marriage is a significant driver of low fertility. Establishing the causal relationship between late marriage and low fertility is challenging, however, because age at first marriage is confounded with many factors. While the most powerful tool that can be used to establish the causality between the two factors is a well-designed randomized controlled trial, implementing such a trial is difficult for ethical reasons.

In this paper, we exploit a quasi-random variation observed for age at first marriage induced by Korea's school entry policy and estimate the effect of age at first marriage on fertility. Specifically, we use a female's exact date of birth as an instrument for age at first marriage, because date of birth determines school age. According to Article 15 of the Enforcement Decree of the Elementary and Secondary Education Act, children who reach school age on March 1 or after must enter elementary school. Because of Article 15, those born between February 1 and February 28 enter elementary school about a year earlier than those born between March 1 and March 31. This school entry policy has large effects on the timing of the completion of education, and because females in Korea usually marry after completing their education, this policy also has large effects on age at first marriage. Date of birth is likely to be exogenous, and we use this variation to control for the endogeneity of age at first marriage. The research design we use is a fuzzy regression discontinuity design (RDD) because date of birth is suitable as a running variable required for the RDD.

We reach three key conclusions from the statistical analysis: First, there is a large and statistically significant impact of date of birth on age at first marriage. Specifically, the estimated difference in age at first marriage between those born in February and March is about 1.5 years. Based on the fuzzy RDD analysis, we find that a delay of one year in age at first marriage reduces the likelihood of a female having any children (extensive margin) by 8 to 10 percentage points, which is statistically significant. Second, this delay leads to a reduction in total childbirths (intensive margin) by 0.1 to 0.15 children, though the estimates are slightly less significant (p value = 0.15). Third, age at first marriage also affects age at first childbirth. A one-year delay in marriage leads to a six-month delay in age at first childbirth. The estimated effects are, however, estimated imprecisely, depending on the bandwidth choice of the RDD, so we caution against deriving a definitive conclusion on this outcome variable.

This paper is structured as follows: We first review prior research relevant for this study. Next, we provide the institutional background, followed by empirical strategies and data description. Then, we show the statistical results and examine the robustness of our findings. Lastly, we derive some policy implications and conclude.

Literature Review

Theoretically, five factors are assumed to affect fertility. The first is the medical factor. Many disorders that induce sterility are prevalent these days, and these lead to low fertility. The second one is related to ideology. Self-realization and individualism have become more prevalent among people who tend to put more emphasis

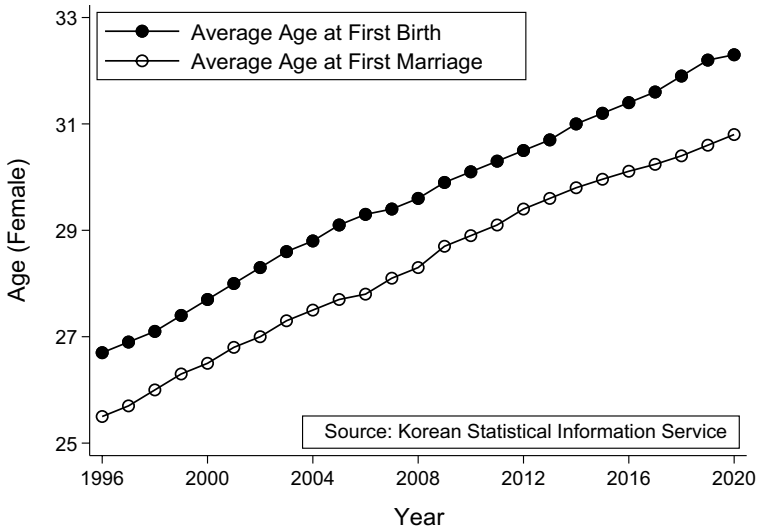
on children's quality than quantity (Morgan, 2003). Third, the role of women has changed significantly. More females are participating in labor markets than previously, raising the opportunity costs of childbirth, and this in turn has led to a low birth rate (Becker, 1991). The fourth factor is high childcare expenses. The level of education that children receive these days has increased significantly, but at the same time, educational expenses in high-income countries have not decreased. Accordingly, parents cannot afford to raise many children (Raymo et al., 2015). Fifth, a delay in age at first marriage reduces the probability of having many children (Jones, 2007). In 1990, the mean age at first marriage across OECD countries was 24.9 for women. By 2014, the mean age has increased to 31.

Of these five factors, this study analyzes the causal impact of age at first marriage on fertility using quasi-experimental research designs. Therefore, we review only the research that pertains to age at first marriage. A number of studies have examined the factors that affect age at first marriage. Ikamari (2005) and Isen and Stevenson (2008), for example, analyzed whether females' education affects age at first marriage. Higuchi (2001) and Gutiérrez-Domènech (2008), by contrast, examined the effect of labor market participation on marriage age. All the studies used age at first marriage as an outcome variable and investigated whether there were any factors influencing this outcome. Our study differs from these studies in that our study uses age at first marriage as an explanatory variable. Nonetheless, the studies mentioned above are informative because they suggest that age at first marriage is an endogenous variable affected by socio-economic factors such as education, and that researchers need to account for the endogeneity issue inherent in the age variable.

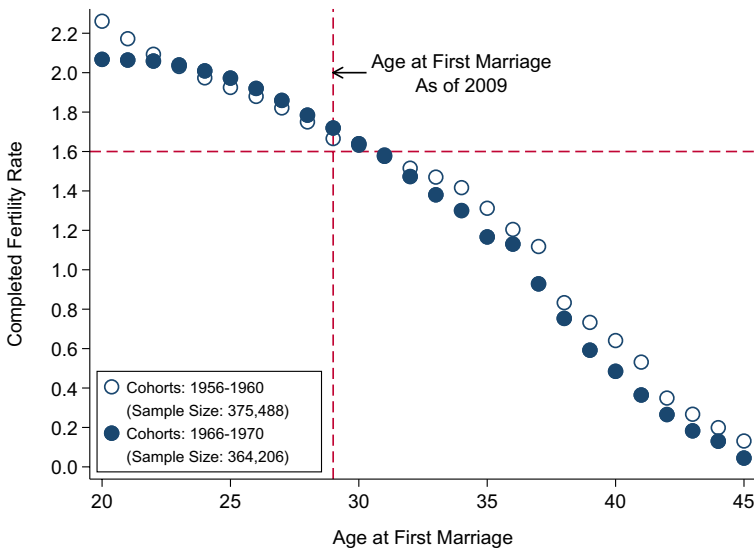
Theoretically, age at first marriage can have significant impact on fertility-related outcomes through many channels. Many studies show that childbearing ages are between ages 15 to 50, and the probability of pregnancy declines as females become older (Dunson et al., 2004). Accordingly, too much delay in marriage timing would make some females to forego childbearing for biological reasons. On the other hand, it is possible that if females delay marriage, however, some females may catch up with those who married early in terms of fertility-related outcomes. Even if this is the case, however, delay in marriage may lead to a lower number of childbirths due to a reduced reproductive life span.

A negative effect of delay in marriage age on fertility is likely to be salient in Korea because age at first marriage is around 29 during the analysis period (i.e., 2009) and is continually increasing (see Panel A of Fig. 1). American Society for Reproductive Medicine shows that the probability of pregnancy for a healthy female at age 30 is only 20%. Thus, it is highly likely that further delay in marriage timing in the context of Korea would result in less childbearing. In Panel A of Fig. 1, it is also evident that correlation between age at first marriage and age at first birth is extremely high. Panel B of Fig. 1 shows the correlation between age at first marriage and completed fertility rate. As can be seen from the figure, age at first marriage is negatively associated with completed fertility rate.¹ The observed simple correlation

¹ Statistics on completed fertility rate is estimated from the 2015 2% census data released by the Statistics Korea.



Panel A. Age at First Marriage and Age at First Birth



Panel B. Age at First Marriage vs. Completed Fertility Rate

Fig. 1 Marriage timing and fertility tempo and quantum

between marriage timing and fertility tempo and quantum suggests that marriage timing is likely to affect fertility-related events in Korea.

A few studies have analyzed the effect of age at first marriage on fertility. Marini (1981)'s study was the first to examine the relationship between the age variable and

fertility. Using a 1973 follow-up survey of high school cohorts in 1957, she finds that the timing of marriage is associated negatively with females' first birth. Some researchers have studied cases in Bangladesh (Kabir et al., 2001; Bates et al., 2007; Nahar et al., 2013). Using the 2002 data, Bates et al. (2007) find that there is a negative correlation between age at first marriage and fertility, although this correlation disappears when they control for other variables. Nahar et al. (2013) find that a year's delay in age at first marriage is associated with 0.7 years in age at first birth. They also find that the delay reduces total births by 0.2 children. Finally, Raj et al. (2009) analyze the India data and find that age at first marriage is correlated with many fertility-related factors.

We can derive several conclusions from prior research. First, the estimated effects of age at first marriage on fertility are mixed. We argue that the two reasons can account for such inconsistency. The first reason is the scope of the studies. The effect of the age variable on fertility is likely to be different depending on the context of the countries and time, which is why we observe varying effects. The second reason is the identification strategy used in previous research. All the studies mentioned above rely on the control function method. While the control function method may control for observable factors that are correlated with fertility and age at first marriage, it is limited in controlling for unobservable factors that are also correlated with the two variables. It is quite likely that the estimated effects found in existing studies reflect both the effect of marriage timing and unobservable variables that affect fertility. For example, if females who married late are also those who care more about their career, the negative association between marriage age and fertility may have been driven by the age, other unobservable factors such as career aspiration, or both. Thus, the estimated effects in previous research may have been driven completely by unobservable factors that affect fertility rather than the age itself. In this study, we exploit the exact date of birth that determines the timing of school entry and use this plausibly exogenous variation as an instrument to isolate the effect of age at first marriage and fertility outcomes. Many studies have used similar instrument as our study such as Black et al. (2011), Kamb and Tamm (2022), McCrary and Royer (2011), and Nguyen and Lewis (2020).

Black et al. (2011), for example, analyze a causal impact of school starting age on educational attainment. Because school starting age is endogenous, the authors exploit exogenous variation in month of birth and use such variation as an instrument for school starting age. Similarly, Nguyen and Lewis (2020) use a month and year of birth as an instrument for school starting age to identify the effect on teenage marriage and motherhood. Interestingly, the authors find the direct impact of school starting age on teenage marriage and motherhood. McCrary and Royer (2011), on the other hand, use an exact date of birth as an instrument for female education to establish causal relationship between female education and fertility and infant health.

Second, most of the previous literature has concentrated on low-income countries. In these countries, the issue is not related to late marriage. Rather, child marriage is the problem. In Bangladesh and India, for example, the mean age at first marriage is 15 and 18, respectively. Consequently, the studies that examine these countries are concerned with child marriage and high fertility. The conclusions of

these studies, therefore, may not be applicable to high-income countries that are suffering from late marriage and low fertility.

All things considered, our study contributes to the existing literature mainly in two ways. First, we argue that the internal validity of the estimated effect of age at first marriage on fertility is relatively stronger than that of other studies because this study exploits a quasi-random variation in marriage timing. Specifically, we exploit exact date of birth as an instrument for females' age at first marriage and use a fuzzy RDD to estimate the causal impact of age at first marriage on fertility. Second, the variation observed for age at first marriage in this study is at the age of 27. Although the mean age at first marriage in most countries facing low fertility varies depending on the surveyed year, the mean age ranges from 24 to 32 in these countries.² We, therefore, believe that this study provides helpful policy implications for such countries.

Institutional Background

In this paper, we use the South Korean context to analyze the causal impact of age at first marriage on female fertility. Analyzing the Korean context helps validate the statistical relationship between marriage timing and fertility because Korea is unique in the sense that the dynamics of marriage timing and fertility is continuously changing. To analyze the effect of marriage timing and fertility, it must be the case that the two variables must covary to some extent. Without such covariance, it is difficult to identify a causal relationship between the two variables. Because the two variables are continuously changing, analyzing the Korean context allows us to identify whether causality exists between the two factors.

The validity of this study hinges critically on variation observed for age at first marriage and on such variation being exogenous. The exogenous variation we exploit is the school entry policy in Korea. According to Article 15 (Making List of Schoolchildren) of the Enforcement Decree of the Elementary and Secondary Education Act, those who reach elementary school age on March 1 of year t are admitted to elementary school in year t , but those born on February 28 or before of year t are admitted to elementary school in year $t - 1$. Therefore, children born in February enter elementary school a year ahead of those born in March of the same birth year. For example, suppose there are two children, A and B. Child A is born in February 2000, and Child B is born in March 2000. Child A will be admitted to elementary school in March 2006, whereas Child B will be admitted to elementary school in March 2007, even though both children were born in 2000.

Note that Article 15 was amended by Presidential Decree on May 27 of 2008. According to the amended article, the two children (A and B) are admitted to elementary school in the same school year, and so we no longer observe variation in school entry dates. Moreover, parents can now choose the timing of school entry

² In 2015, the mean age was 24, 29, 27, 32, and 32, for China, Japan, the United States., Spain, and Germany, respectively.

based on the amended article. Consequently, it is less likely that the variation observed for school entry dates of birth cohorts exposed to the amended article are exogenous. The birth cohorts used in our analysis were not affected by the amended article above; therefore, the variation in school entry dates is plausibly exogenous.

The variation in school entry year is likely to affect age at first marriage of Korean females for the following reasons. First, elementary and middle school is mandatory in Korea. Consequently, a year difference in school entry results in a year difference in middle school graduation. Second, a year difference in middle school graduation is likely to induce a year difference in high school graduation. In Korea, the high school enrollment rate among middle school graduates is very high. According to Statistics Korea, the enrollment rate has been close to 90% since 2000. Furthermore, dropout and retention rates in high school are very low (less than 5%), and more than 90% of students graduate high school on time. In general, therefore, if Child A was admitted to elementary school a year before Child B, the chance that Child A completes education a year before Child B is very high. Third, the legal marriage age in Korea is 18 (i.e., after graduating high school). According to the 2006 vital statistics survey, only 2.3% of females marry before the age of 19. This shows that, in Korea, females typically marry after finishing their high school education. Putting these facts together, we can conclude that it is quite likely that Child A will enter the marriage market a year before Child B because of the school entry policy in Korea.

We, therefore, exploit the institutional background mentioned above as a plausibly exogenous variation in age at first marriage and isolate the causal impact of age at first marriage on fertility. Specifically, we compare those born in February of birth year t with those born in March of the same birth year and examine whether the latter cohorts marry later than the former. If we see a difference in age at first marriage between the two groups, we use this variation to estimate the effect of age at first marriage on fertility.

Empirical Strategy and Data

Empirical Strategy

This study uses the fuzzy RDD to analyze the causal relationship between age at first marriage and fertility. To estimate the effect (τ_i) of age at first marriage on fertility in the context of the fuzzy RDD, researchers need to estimate the following four conditional expectation functions:

$$\tau_i = \frac{\lim_{x \rightarrow 0^+} E(Y_i | X_i = x) - \lim_{x \rightarrow 0^-} E(Y_i | X_i = x)}{\lim_{x \rightarrow 0^+} E(D_i | X_i = x) - \lim_{x \rightarrow 0^-} E(D_i | X_i = x)}$$

Here, Y_i , X_i and D_i are the outcome variable (fertility), running variable (date of birth), and treatment variable (age at first marriage), respectively, with the fact that the running variable predicts the treatment variable as follows:

$$0 < \lim_{x \rightarrow c^+} P(D_i = 1 | X_i = x) - \lim_{x \rightarrow c^-} P(D_i = 1 | X_i = x) < 1$$

Here, c denotes the cutoff point for the running variable. The cutoff is the distance (in days) between a female’s date of birth and March 1 in birth year t . Suppose female i is born on February 27 in year t , and female j is born on March 1 in year t . Then the values of the running variable for female i and j are -2 and 0 . As mentioned previously, females who are born on February enter elementary schools about a year before those who are born on March. Because whether a female is born on February 28 or March 1 is exogenous, school starting age is plausibly exogenous for females born around this cutoff. We, therefore, use this fact as an instrument for age at first marriage. Exploiting this setting, the equation above implies that the probability of marrying late (i.e., $D_i = 1$) is higher for those born on March 1 or after (i.e., $c \geq 0$) than for those born on February 28 or before (i.e., $c < 0$).³

To estimate the four conditional functions, the statistics literature proposes using two estimators: local polynomial or global polynomial regression estimators. While the two estimators have their own advantages and disadvantages, the RDD literature recommends using local polynomial estimators because they have desirable properties in the context of RDD (Hahn et al., 2001; Imbens & Lemieux, 2008; Lee & Lemieux, 2010). To estimate the local polynomial regression estimators, researchers need to solve the following four minimization problems:

$$\min_{\alpha_r, \beta_r^p} \sum_{0 \leq X_i < +h} (Y_i - \alpha_r - \beta_r^1 X_i - \beta_r^2 X_i^2 - \dots - \beta_r^p X_i^p) K\left(\frac{X_i - x}{h}\right)$$

$$\min_{\alpha_l, \beta_l^p} \sum_{-h \leq X_i < 0} (Y_i - \alpha_l - \beta_l^1 X_i - \beta_l^2 X_i^2 - \dots - \beta_l^p X_i^p) K\left(\frac{X_i - x}{h}\right)$$

$$\min_{\alpha_r, \beta_r^p} \sum_{0 \leq X_i < +h} (D_i - \alpha_r - \beta_r^1 X_i - \beta_r^2 X_i^2 - \dots - \beta_r^p X_i^p) K\left(\frac{X_i - x}{h}\right)$$

$$\min_{\alpha_l, \beta_l^p} \sum_{-h \leq X_i < 0} (D_i - \alpha_l - \beta_l^1 X_i - \beta_l^2 X_i^2 - \dots - \beta_l^p X_i^p) K\left(\frac{X_i - x}{h}\right)$$

The first two minimization problems estimate the first two conditional expectation functions, while the last two minimization problems estimate the last two conditional expectation functions.

When solving the minimization problems, researchers have to make choices for three parameters. The first parameter is p , which determines the degree of polynomial. If $p = 1$, then we are estimating the treatment effects using a local linear

³ For those whose exact date of birth is February 29, we coded them as February 28 for the sake of the statistical analyses.

regression specification. If $p = 2$, then we are estimating the effects using a local quadratic regression specification. For the sake of transparency of research, the RDD literature proposes showing the estimated effects using several polynomial specifications, rather than resorting to a single polynomial specification. In addition, a recent study recommends that researchers estimate the effect using $p = 1$ or $p = 2$ (Gelman & Imbens, 2019). We, therefore, present estimated effects obtained from local linear and local quadratic specifications. Another parameter that researchers have to choose is K , the kernel function. This determines the relative weight that each observation receives in an estimation process. While many types of kernel function have been developed to date, the RDD literature typically uses uniform or triangle kernel functions. We, therefore, provide effect estimates derived from both functions.

The third parameter is the bandwidth choice, arguably the most important choice that researchers have to make in RDD. The validity of the RDD relies on the observations to the left (control group) and right (treated group) of the cutoff point of a running variable being comparable. Because the two groups are similar, any difference in an outcome variable can be attributed to the effect of a treatment. It is likely, however, that observations located at the left end of a running variable are not comparable to those at the right end of a running variable.⁴ In the RDD, therefore, the effect estimate is derived from using observations near the cutoff point; the critical issue is determining how near the observations should be (i.e., the bandwidth choice). The RDD literature proposes using several bandwidth choices and seeing whether the estimated effects are sensitive to the choice of bandwidth. If the effect estimates are too sensitive to the choice of bandwidth, it is doubtful that the effects are causally estimated. Following this recommendation, we present effect estimates derived from several bandwidth choices.

A local polynomial regression estimator used in an RDD is basically a weighted ordinary least squares because it derives a point estimate by giving different weights to each observation. Furthermore, the fuzzy RDD used in this study can be thought of as a weighted instrumental variable estimator. Thus, in order to conduct statistical inference properly, researchers have to account for the heteroscedasticity issue when estimating standard errors. In an RDD, observations are typically clustered at the level of a running variable, and hence researchers should account for the clustering issue. Lee and Card (2008) recommend clustering the standard errors at the level of a running variable when conducting a local polynomial regression estimator. We, therefore, conduct statistical inference using cluster-robust standard errors clustered at the level of date of birth.

Data

In this study, we use the 2009 Nationwide Fertility Level and Family Health and Welfare Survey, administered jointly by the Ministry of Welfare & Family and the

⁴ If we think of a running variable as household income, then the statement is clear.

Korea Institute for Health and Social Affairs. We use the 2009 data rather than other years because 2009 is the only year that contains information on exact date of birth. If this variable does not exist in the dataset, we cannot exploit an exogenous variation created by the school entry policy. While the exact date of birth is not publicly available, we were able to obtain them by following the formal data application process.⁵

The data contain not only exact date of birth but also other useful variables that can be used for implementing the RDD. For example, the dataset contains information on age at first marriage, so we can statistically test whether the variation in age at first marriage is induced by exact date of birth. The data also include information on outcome variables such as the total number of births and age at first birth, as well as some baseline covariates that can be used to test the continuity in observable characteristics at the cutoff point of the running variable between the control and treated group, an assumption critical for the validity of the RDD. When analyzing the effect on fertility, we used two outcome variables: any childbirth and total number of childbirth. The former corresponds to the extensive margin, whereas the latter indicates the intensive margin. Each margin is likely to be affected by different factors, so analyzing which margin is more pervasive allows us to develop more efficient policy measures (Feyrer et al., 2008).

The birth cohorts used for the analysis are those born between 1973 and 1982. The first reason we do not include those who were born after 1982 is that there are very few females from this cohort who are married. In 2009, the mean age at first marriage in Korea was 28.7. The age of females who were born in 1983 was 27 in 2009. Therefore, because the data is representative of the Korean population, it is natural that the data do not contain married women who were born in 1983 or after. The second reason is that few of these females have experienced childbirth. In Korea, the share of extramarital childbirth is less than 2%. That is why we do not observe any births for these females.⁶ There are two reasons for using those who were born in 1973 or after. First, Korea experienced a disastrous economic crisis in 1998. We believe that the birth cohorts who experienced the economic crisis at the time of marriage are not comparable to those who did not experience the crisis. The age of those who were born in 1973 was 25 in 1998, and the mean age at first marriage was around 25 in 1998. We, therefore, decided to use the 1973 birth cohorts as the beginning of the analysis sample. Second, it is likely that the factors that determine fertility are not the same across generations. Consequently, we argue that including many birth cohorts in the analysis sample is not appropriate for generalizing the conclusion of the study.

⁵ We thank the Korea Institute for Health and Social Affairs for generously providing the information on exact date of birth.

⁶ Note that because the share of childbirths outside of marriage is extremely small, the analysis sample does not include females who gave childbirths before getting married.

Results

When implementing RDD, we follow six steps.

- a. Tests of variation in a treatment variable: we test for discontinuity in age at first marriage at the 0 vertical line (i.e., the cutoff point of a running variable that corresponds to March 1). If we do not observe any statistically significant discontinuity in our treatment variable that is large in magnitude, then it is questionable whether the observed discontinuity in an outcome variable, if any, can be attributed to the effect of the treatment.
- b. Tests of manipulation: we test for discontinuity in the density of a running variable. The RDD requires that people have imprecise control over a running variable. Thus, we should not observe any statistically significant discontinuity in the density of a running variable.
- c. Tests of balance in baseline characteristics: another identification assumption required for the validity of the RDD is that the two groups are similar. Accordingly, we should not observe any statistically significant discontinuity in baseline covariates.
- d. Examining the validity of exclusion restriction: the most important assumption in this study is that the instrument affects fertility “only” through marriage age (i.e., the instrument does not have a direct influence on fertility). While proving directly the validity of exclusion restriction assumption is not possible, we present several arguments to support the hypothesis that the main channel through which the instrument affects fertility-related events is marriage age.
- e. Tests of discontinuity in outcome variables: if there is a causal relationship between age at first marriage and fertility, we should observe statistically significant discontinuity in an outcome variable.
- f. Falsification tests: if we observe a discontinuity in fertility at March 1 and were to attribute such a discontinuity to the difference in age at first marriage, we should not observe any statistically significant discontinuity at other cutoff points, because the school entry policy is irrelevant at other cutoff points. In order to promote the validity of our conclusions, therefore, we conduct some falsification tests by estimating the discontinuity at other cutoff points, such as April 1.

In all of the steps above, we first present graphical analyses followed by local polynomial regression results.

Tests of Variation in Age at First Marriage

The validity of this study is dependent critically on the fact that we observe statistically and significant discontinuity in females’ age at first marriage at the running variable’s cutoff point. Our running variable is date of birth, and the cutoff is March 1. In Fig. 2, we present density of age at first marriage (representing the y -axis) by the running variable (representing the x -axis). The density of the figure is created

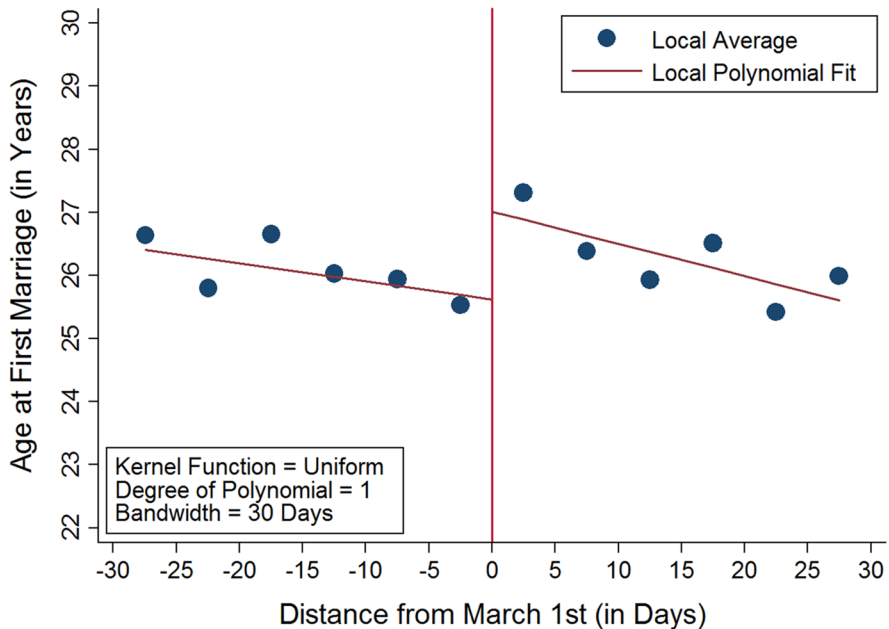


Fig. 2 Density of age at first marriage by the running variable

using the uniform kernel function with a binwidth of 5 days and bandwidth of 30 days. The local polynomial fit is based on a local linear specification.⁷ Each dot represents the mean age at first marriage for the corresponding value of the running variable. For example, the dot within the range of -30 and -26 indicates the estimated mean age at first marriage obtained from averaging the values of observations whose exact date of birth is between January 30 and February 3.

As can be seen from the figure, age at first marriage for those born before March 1 (left of the cutoff) is around 26, while the age for those born on March 1 or after (right of the cutoff) is around 27. Moreover, we see a discontinuity at the zero vertical line that is visually clear. Although we cannot determine the exact magnitude of the discontinuity and its statistical significance by looking at the figure, the discontinuity we observe at the cutoff is about one year. To put this in context, those whose exact date of birth was March 1 or after married about a year later than those whose exact date of birth was before March 1. This finding is consistent with the variation we exploited in this study.

While examining the variation in age at first marriage graphically is important and helpful in the context of the RDD, it does not allow us to test this variation statistically. Another limitation of the graphical result in Fig. 2 is that the two groups (left and right) are compared without accounting for the birth year.

⁷ For the sake of consistency, we use the same specifications for graphical analyses throughout the paper. Note that the graphical results do not change qualitatively even if other specifications are used.

Table 1 Tests of Discontinuity in the Treatment Variable (Age at First Marriage)

Bandwidth choice	Local linear model		Local quadratic model	
	Uniform	Triangle	Uniform	Triangle
20 days	1.493*** (0.493)	1.941*** (0.397)	2.327*** (0.468)	2.039*** (0.422)
30 days	1.283*** (0.451)	1.594*** (0.412)	2.028*** (0.479)	2.129*** (0.403)
40 days	0.938** (0.473)	1.321*** (0.420)	1.800*** (0.485)	1.969*** (0.430)
Birth year fixed effects	Yes			
Number of observations	1876			

Standard errors—clustered at the distance level—are in parentheses. Birth years used in the analysis are those who are born from 1973 to 1982. “Uniform” and “Triangle” correspond to the kernel function used in the estimation. Age at first marriage—the treatment variable—in years. The bandwidth choice of 20 days indicates the distance (in days) from March 1st

*** and ** indicate statistical significance at the 1 and 10% level

In order to analyze the causal impact of age at first marriage on other variables, we must compare the two groups within the same birth year. Hence, in Table 1, we present a formal test of the discontinuity in age at first marriage using local polynomial regression. As mentioned in the Empirical Strategy and Data section, we test the discontinuity at the cutoff under various specifications (i.e., p , K , and h). And importantly, all the regressions are conditional on birth-year fixed effects so that the comparison is based on the same birth cohort. Standard errors are clustered at the distance level.

The first two columns show discontinuity estimates based on a local linear specification; the first column uses uniform kernel function, and the second column uses triangle kernel function. Under the bandwidth choice of 20, 30, and 40 days and uniform kernel function, the estimated discontinuity in age at first marriage is 1.493, 1.283, and 0.938, respectively, indicating that those who were born on March 1 or after married about 1 to 1.5 years later than did those who were born before March 1. The last two columns show the same estimates based on a local quadratic specification. The discontinuity estimates become a little bit larger under a local quadratic specification: a difference of two years. While the exact estimates change a little bit depending on specifications, the estimates are quite similar qualitatively. All the estimates show that those who entered elementary school a year later married about one to two years later than did those who entered earlier. In addition, all the estimates are statistically significant. In sum, we see a statistically significant discontinuity that is quite large in magnitude in females’ age at first marriage at the cutoff point. We use this variation to isolate the causal impact of age at first marriage on fertility.

Tests of Manipulation in the Running Variable

When using an RDD, it is important that the variation observed in a treatment variable is exogenous. A test that was developed for use in testing such exogeneity is the density test (McCrary, 2008). The rationale behind this test is that if people can manipulate the value of a running variable (e.g., date of birth) in an effort to receive a treatment (e.g., child entering school early), we would observe a statistically significant discontinuity in the density of the running variable at the cutoff point. Such a discontinuity would suggest that people to the left and to the right of the cutoff point are not comparable. We show the density of the running variable in Fig. 3 (top). The density of the running variable is smooth across the values of the running variable, and we observe almost no significant discontinuity at the zero cutoff. To gauge the statistical significance of the discontinuity by bandwidth choice, we show in Fig. 3 (down) the discontinuity estimates under many bandwidth choices. As can be seen from the figure, all the estimates are close to zero, with the 95% confidence intervals enclosing the zero horizontal line, implying that the estimates are insignificant at the 5% level. The results of the density test, therefore, suggest that manipulation is unlikely in this context.

While we do not observe any manipulation in terms of statistical analyses, it is necessary for researchers to examine the practical likelihood of such manipulation. In Korea, it is almost impossible for parents to manipulate their child's date of birth, because parents have to submit a certificate of birth provided by a hospital when registering the birth of their child. In addition, we argue that even if such manipulation were possible, such as by lobbying doctors at a hospital, there are few incentives for parents to engage in such manipulation in order to enter school early. Besides, if the second reason is pervasive, we should definitely observe some discontinuities at the cutoff, which is not the case here. We argue, therefore, that manipulation is not an issue in this study.

Tests of Balance in Baseline Covariates

Another identifying assumption in an RDD is that observations left and right of a cutoff point are similar in terms of observable and unobservable characteristics. Just as in a randomized controlled trial, in which we test for the balance in baseline characteristics between a treated and controlled group, we should test for discontinuity in such characteristics at the cutoff point. In Fig. 4, we provide densities of baseline covariates by the running variable. The first two figures at the top plot densities of the share of college graduates and whether a female's husband is a college graduate. While there appears to be some variation in the share across the running variable, we do not observe any significant discontinuity at the cutoff. Our data indicate that the mean share of college graduates for females and husbands is about 40%. The two figures at the center correspond to the densities of the share of females who are employed right before their first marriage and the share of females who are employed as regular employees. As can be seen from the two figures, the mean

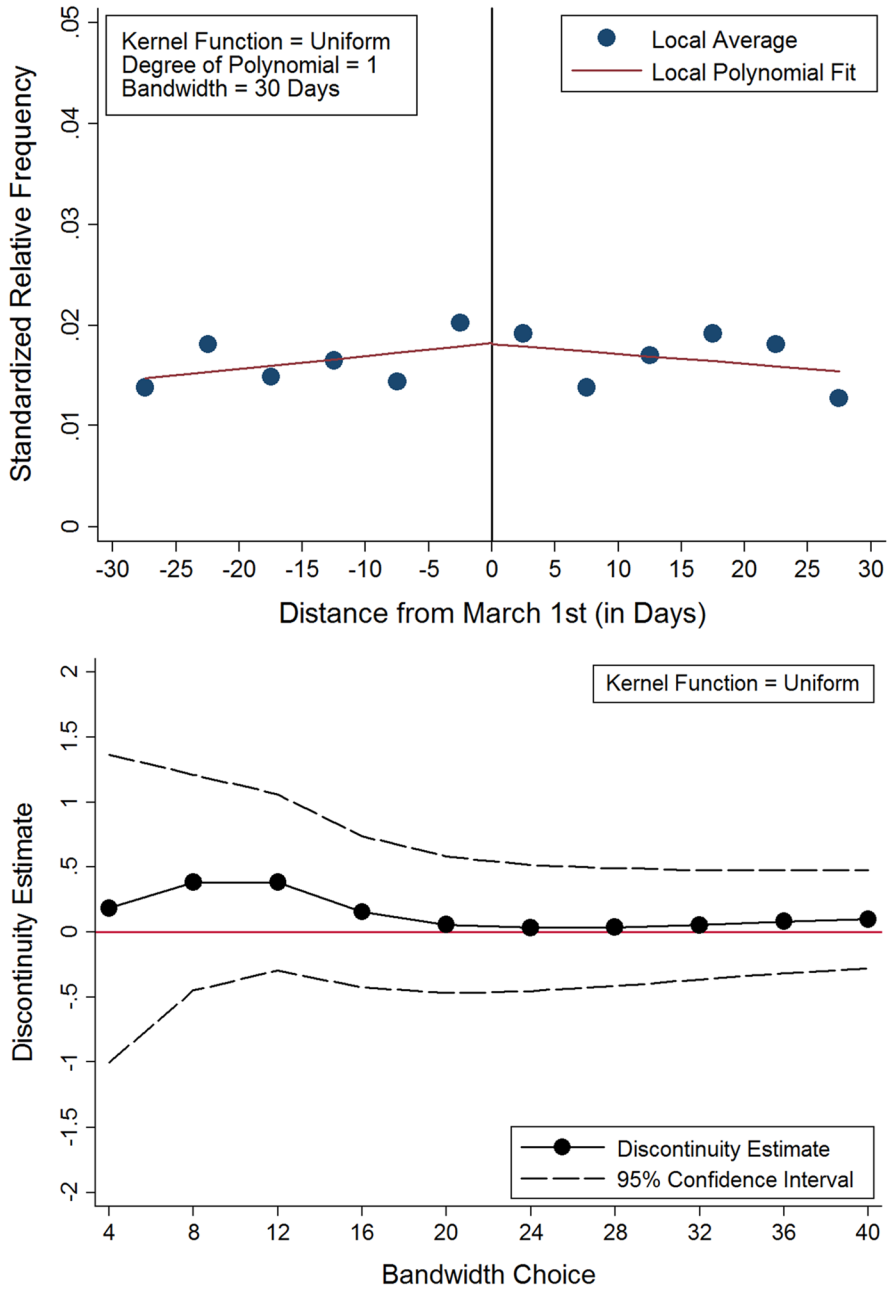


Fig. 3 Density of the running variable (top) and discontinuity estimates (bottom)

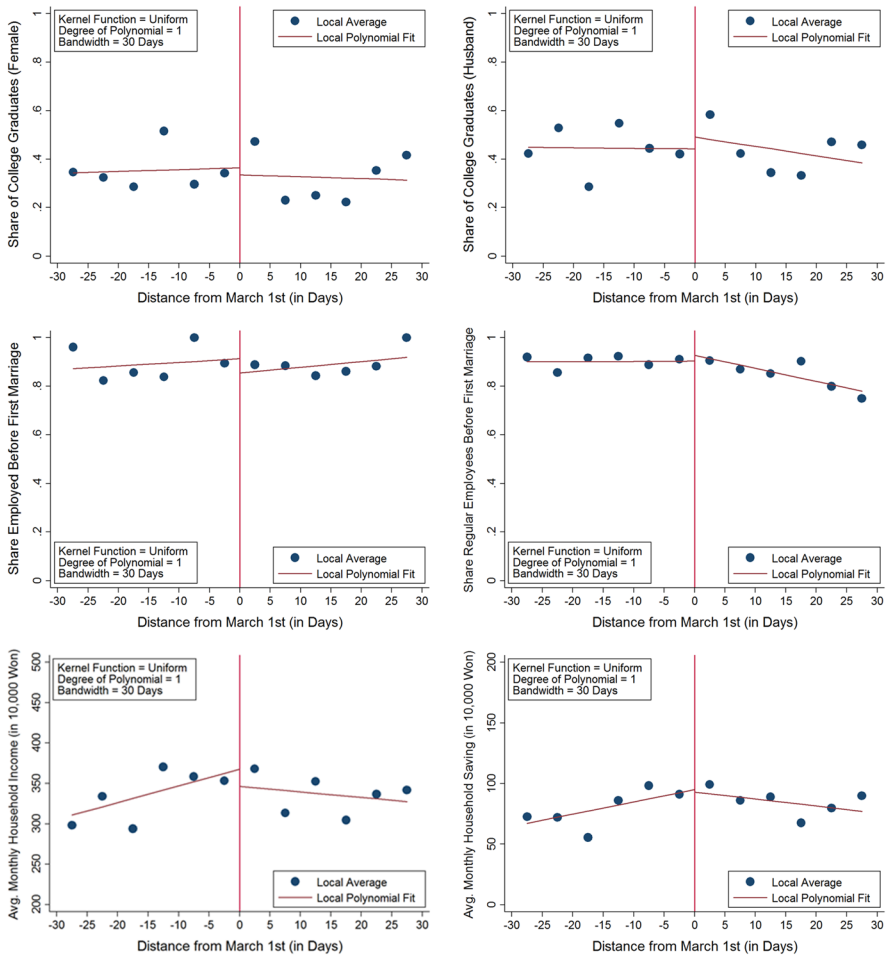


Fig. 4 Density of baseline covariates by the running variable

share for each variable is about 90%, implying that most females were employed as a regular employee right before they first married. In addition, we do not observe any discernible discontinuities at the cutoff.

The last two figures at the bottom of Fig. 4 indicate the financial conditions of households. For the financial conditions, we analyzed household income and savings. Note that these two variables are not determined prior to our treatment variable. The financial conditions are surveyed at the time of this survey data (i.e., 2009). Thus, while observing discontinuities in these two variables does not necessarily imply that the identifying assumption is violated, we test for the continuity in these variables in order to examine whether the two groups are substantially different financially at the time of survey. As can be seen from the last two figures, we do not

Table 2 Tests of discontinuity in baseline covariates (local linear regression estimates)

Baseline covariate	Bandwidth choice		
	20 Days	30 Days	40 Days
Panel A: Uniform Kernel Function			
Finished college (wife)	0.065 (0.115)	- 0.035 (0.107)	- 0.051 (0.094)
Finished college (husband)	0.072 (0.094)	0.040 (0.087)	0.033 (0.079)
Was employed before marriage	- 0.062 (0.075)	- 0.058 (0.062)	- 0.037 (0.052)
Was worked as regular employees	0.009 (0.057)	0.020 (0.055)	- 0.004 (0.049)
Household income	- 27.318 (40.742)	- 30.868 (34.731)	- 16.138 (30.197)
Household saving	- 10.612 (17.717)	- 5.845 (15.403)	3.516 (13.073)
Panel B: Triangle Kernel Function			
Finished college (wife)	0.123 (0.126)	0.043 (0.103)	- 0.015 (0.097)
Finished college (husband)	0.137 (0.098)	0.073 (0.085)	0.043 (0.080)
Was employed before marriage	- 0.025 (0.085)	- 0.058 (0.071)	- 0.045 (0.061)
Was worked as regular employees	0.019 (0.059)	0.014 (0.053)	0.009 (0.048)
Household income	0.328 (44.385)	- 20.204 (36.720)	- 21.406 (32.268)
Household saving	- 1.545 (19.136)	- 6.935 (16.312)	- 2.937 (14.074)
Birth cohort fixed effects	Yes		
Number of observations	1876		

Standard errors—clustered at the distance level—are in parentheses. Birth cohorts used in the analysis are those who are born from 1973 to 1982. For the “Was employed as office workers” and “Was worked as regular employees” variables, the number of observations used in the estimation is 1702. The bandwidth choice of 20 days indicates the distance (in days) from March 1st. The variables “Finished college,” “Was employed before marriage,” “Was employed as office workers,” and “Was worked as regular employees” are dummy variables. The variables “Household income,” “Household consumption,” and “Household saving” are in 10,000 Korean won (10,000 won is approximately 10 dollars)

observe any salient discontinuity at the cutoff point, implying that the two groups are similar in terms of financial conditions.

Finally, we provide in Tables 2 and 3 the results of the tests of discontinuity in baseline characteristics under various parameter choices. Table 2 presents results of local linear regression estimators. None of the discontinuity estimates is statistically

Table 3 Tests of Discontinuity in Baseline Covariates (Local Quadratic Regression Estimates)

Baseline covariate	Bandwidth choice		
	20 Days	30 Days	40 Days
Panel A: Uniform Kernel Function			
Finished college (wife)	0.213 (0.191)	0.157 (0.141)	0.038 (0.125)
Finished college (husband)	0.232* (0.140)	0.130 (0.103)	0.067 (0.100)
Was employed before marriage	0.007 (0.113)	-0.068 (0.095)	-0.067 (0.083)
Was worked as regular employees	0.025 (0.073)	0.015 (0.072)	0.033 (0.068)
Household income	51.441 (55.829)	7.132 (46.854)	-21.881 (41.147)
Household saving	19.940 (22.694)	-0.789 (20.689)	-7.983 (18.385)
Panel B: Triangle Kernel Function			
Finished college (wife)	0.052 (0.165)	0.160 (0.155)	-0.015 (0.097)
Finished college (husband)	0.126 (0.132)	0.166 (0.114)	0.043 (0.080)
Was employed before marriage	0.004 (0.113)	-0.020 (0.100)	-0.045 (0.061)
Was worked as regular employees	0.021 (0.071)	0.024 (0.067)	0.026 (0.063)
Household income	41.501 (50.315)	21.438 (50.473)	-4.064 (44.303)
Household saving	16.837 (21.043)	7.526 (20.716)	-3.687 (19.096)
Birth cohort fixed effects	Yes		
Number of observations	1876		

Standard errors—clustered at the distance level—are in parentheses. Birth cohorts used in the analysis are those who are born from 1973 to 1982. For the “Was employed as office workers” and “Was worked as regular employees” variables, the number of observations used in the estimation is 1702. The bandwidth choice of 20 days indicates the distance (in days) from March 1st. The variables “Finished college,” “Was employed before marriage,” “Was employed as office workers,” and “Was worked as regular employees” are dummy variables. The variables “Household income,” “Household consumption,” and “Household saving” are in 10,000 Korean won (10,000 won is approximately 10 dollars)

*Indicates statistical significance at the 10% level

significant, regardless of the bandwidth and kernel function choice. Discontinuity estimates based on local quadratic regression estimators are presented in Table 3. Among 36 discontinuity estimates derived from various parameter choices, only one turned out to be statistically significant at the 10% level. Given the estimated results

Table 4 Share of childbirths outside of marriage in Korea

Variable	Year									
	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
% Childbirths	1.19	1.27	1.40	1.58	1.62	1.79	1.77	1.78	1.79	1.95

in Tables 2 and 3, we argue that the identifying assumption of the RDD is met, and we proceed with the analysis of the effect of age at first marriage on fertility. When analyzing the effect, we control for these baseline covariates for the sake of the efficiency of the regression estimators.

Validity of Exclusion Restriction Assumption

This study uses the March 1st cutoff that creates an exogenous variation in school starting age as an instrument for age at first marriage. One of the critical identifying assumption is the exclusion restriction assumption: the instrument affects fertility-related events “only” through marriage age (i.e., the instrument does not have a “direct” influence on fertility). We present three arguments to support the validity of the exclusion restriction assumption. First, in Korea, fertility-related events are realized only when marriage occurs. Unlike other western countries where the share of childbirths outside of marriage is around 40%, the share in Korea is extremely small (Seo, 2019). According to the official vital statistics, the share of childbirths outside of marriage in Korea is always less than 2% (see Table 4). Because there are few females who give childbirths outside of marriage, it is quite likely that the main channel through which the instrument affects fertility is through marriage-related events.

Second, a large literature that analyzes the effect of school starting age on fertility-related outcomes shows that there are many channels through which school starting age affects fertility (Nguyen & Lewis, 2020). That is, school starting age “indirectly” affects fertility-related events. For example, some research show that age at school entry affects educational level of females, and this mediates the relationship between school starting age and fertility-related events (e.g., McCrary & Royer, 2011). Other research show that health-related variables mediate the relationship between the two variables (e.g., Balestra et al., 2020). In sum, literature that analyzes the effect of school starting age on fertility outcomes is estimating the reduced form effect, and such literature assumes that there are many channels through which school starting age affects fertility-related outcomes. We argue that among various mechanisms, our instrument affects fertility mainly through its effect on marriage timing because of the reason mentioned above.

Third, as the literature revealed, there are many channels through which the instrument indirectly affects fertility such as its effect on educational level. While we cannot test the relationship between our instrument and all the proposed channels, our data allow us to test whether our instrument affects the likelihood of finishing a college. We can also test whether females are more likely to have worked before

marriage and have worked as regular employees. To test whether our instrument affects these three alternative channels, we estimated a regression discontinuity at the cutoff where our instrument turns on. If there are significant discontinuities in these channels, then it calls into the question on whether the main channel through which our instrument affects fertility is marriage timing. The discontinuity estimates are presented in Tables 2 and 3. As can be seen from the tables, we do not see any statistically significant discontinuities at the cutoff point for these channels regardless of the bandwidth and kernel function choices. We argue that the main reason school starting age does not affect the educational level of females is because almost all females graduate high school and about 70% of graduates go on to college. Furthermore, dropout rates are also very small in Korea. According to the Education Statistics of Korea, the dropout rate for middle, high, and university is less than 1, 2, and 4%, respectively.

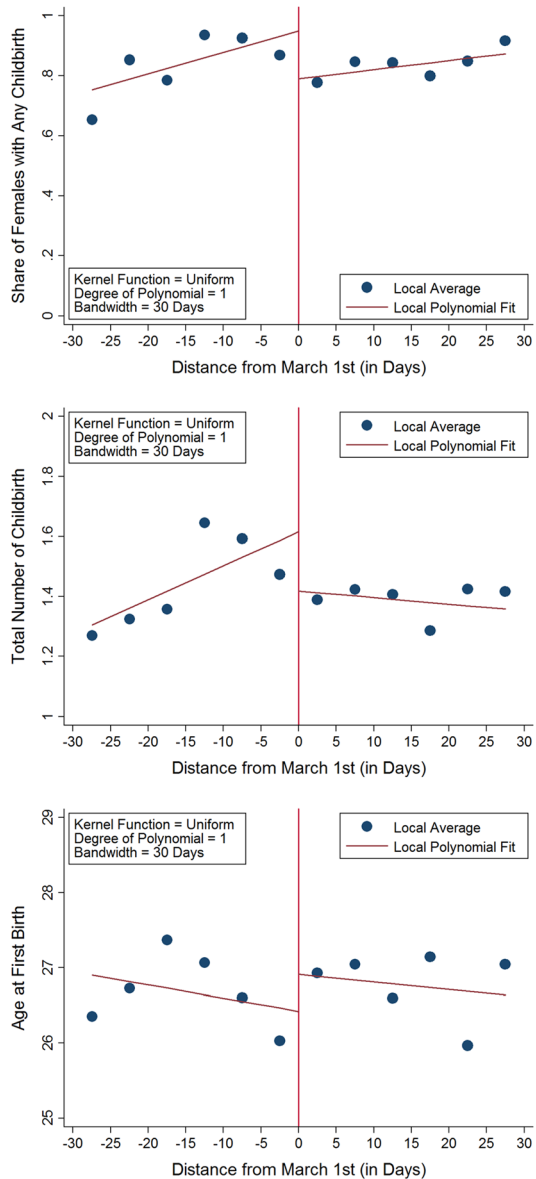
All in all, while the results presented above do not necessarily “prove” that there are no other channels through which our instrument affects fertility other than marriage age, we argue that the arguments above help support the hypothesis that the main channel through which the instrument affects fertility is through marriage age.

Tests of Discontinuity in Outcome Variables

Given that the identifying assumptions of the RDD are satisfied, we conduct a fuzzy RDD to estimate the effect of age at first marriage on fertility-related variables. We first examine whether age at first marriage affects the likelihood of any childbirth. The top figure in Fig. 5 presents the density of the probability of any childbirth by the running variable. Here, the outcome variable is a dummy variable, which is equal to one if a female experienced any childbirth and zero otherwise. In our sample, around 80% experienced childbirth, though we see some variation in the percentage across the running variable. In particular, we see a discernible discontinuity at the zero cutoff; the share of females who experienced childbirth is about 10 percentage points higher for those born before March 1. The discontinuity indicates that the likelihood of experiencing childbirth is higher for those who married a year earlier. The center figure in Fig. 5 displays the density of the total number of child-births by the running variable. The data show that, on average, the total number of child-births is approximately 1.35, which is similar to the estimated TFR observed for Korea in 2009 (i.e., 1.15). For this outcome, we see that there is a discontinuity at the cutoff (about 0.2 children), and the discontinuity is mostly driven by those born between February 13 and 23. The total number of child-births is mostly 1.4 for other cutoff points. The bottom figure in Fig. 5 plots the density of age at first birth. This outcome is examined to identify a possible mechanism that may drive the increase in fertility, if any. The figure shows a discontinuity of 0.5 years at the cutoff point, though it is difficult to draw a definitive conclusion on whether early marriage leads to a decrease in age at first birth due to the noisiness in the density.

The three figures presented in Fig. 5 indicate that a one-year decrease in age at first marriage leads to an increase in the probability of any childbirth and in the total number of child-births, as well as a decrease in age at first birth. Graphical analyses

Fig. 5 Density of outcome variables by the running variable



are, however, limited in determining the statistical significance of an observed discontinuity. In addition, it is necessary to examine the sensitivity of the estimated discontinuity to the choices of the RDD parameters. We, therefore, provide regression results in Table 5 to analyze the statistical and practical significance of the estimated discontinuities at the cutoff point. Panel A in Table 5 presents a discontinuity estimate derived under various parameter choices. While there are some variations in the estimated effects, most of the estimates are around -0.08 , indicating that the

Table 5 Effects of Age at First Marriage on Outcome Variables (Fuzzy RD Estimates)

Specification	Bandwidth Choice		
	20 Days	30 Days	40 Days
Panel A: outcome variable = one or more births			
$k = \text{Uniform}; p = 1$	-0.085* (0.047)	-0.125** (0.058)	-0.094 (0.059)
$k = \text{Triangle}; p = 1$	-0.071** (0.035)	-0.086** (0.041)	-0.100** (0.046)
$k = \text{Uniform}; p = 2$	-0.073** (0.034)	-0.062* (0.037)	-0.112** (0.046)
$k = \text{Triangle}; p = 2$	-0.083*** (0.029)	-0.071** (0.034)	-0.082** (0.036)
Panel B: outcome variable = total number of births			
$k = \text{Uniform}; p = 1$	-0.079 (0.096)	-0.150 (0.096)	-0.166 (0.113)
$k = \text{Triangle}; p = 1$	-0.077 (0.065)	-0.113 (0.080)	-0.135 (0.085)
$k = \text{Uniform}; p = 2$	-0.070 (0.064)	-0.077 (0.077)	-0.132 (0.086)
$k = \text{Triangle}; p = 2$	-0.089* (0.048)	-0.061 (0.064)	-0.092 (0.073)
Panel C: outcome variable = age at first birth			
$k = \text{Uniform}; p = 1$	0.506 (0.517)	0.579 (0.651)	0.185 (1.390)
$k = \text{Triangle}; p = 1$	0.639** (0.284)	0.657** (0.318)	0.552 (0.482)
$k = \text{Uniform}; p = 2$	0.618* (0.385)	0.737*** (0.255)	0.684* (0.371)
$k = \text{Triangle}; p = 2$	0.281 (0.662)	0.533 (0.383)	0.672** (0.310)
Birth cohort fixed effects	Yes		
Number of observations	1876		

Standard errors—clustered at the distance level—are in parentheses. Birth cohorts used in the analysis are those who are born from 1973 to 1982. The bandwidth choice of 20 days indicates the distance (in days) from March 1st. The outcome variable “One or More Births” is an indicator variable. The coefficient estimates correspond to fuzzy regression discontinuity estimates with “Age at First Marriage” being the treatment variable. All the local polynomial regression estimators are conditional on baseline covariates presented in Tables 2 and 3

***, **, and *Indicate statistical significance at the 1, 5, and 10% level

probability that a female experiences childbirth is 8 percentage points lower if a female marries a year later. All but one of the estimates are statistically significant at the 10% level. The control-side mean is about 0.8, so the effect size in percentage terms is about 10%. We can conclude from Panel A that a delay of approximately

five years in marriage leads to a decrease in the likelihood of any childbirth by 40 percentage points.

Panel B presents the estimated effects for the total number of childbirths. The sign of all the estimates is negative, and the magnitude of the estimates is about -0.1 , which coincides with the graphical result presented in Fig. 5 (the figure at the center). The control-side mean is about 1.6 children, so the effect size is about 6.3%. Contrary to the results observed for the likelihood of any childbirth, however, the estimates are statistically insignificant at the 10% level, although most of the estimated p values are less than 0.15. We, therefore, argue that more studies are necessary to draw a definitive conclusion. The last panel in Table 5 shows the estimated effect of age at first marriage on age at first birth. With the exception of two estimates (i.e., 0.185 and 0.281), the estimated effect is on average 0.6 years, implying that an increase of one year in age at first marriage leads to an increase of 0.6 years in age at first birth. It seems that early marriage increases the chance of early childbirth. Note, however, that a half of the estimates are statistically insignificant, so the effects are estimated imprecisely. We, therefore, caution against drawing a strong conclusion on the effect of age at first marriage on age at first birth, and we argue that further studies would help determine the significance of the effect on this outcome.

Falsification Tests

We provide the results of falsification tests to examine the validity of the estimated results presented in Fig. 6 and Table 6. The idea of the falsification tests we conducted is to estimate the discontinuity estimates for other cutoff points. This study exploits the fact that children born in February enter elementary school about a year earlier than do children born in March, and that this variation leads to an exogenous variation in age at first marriage. Using this variation, we find that a one-year increase in age at first marriage leads to a decrease of 8 percentage points in the probability of childbirth. If this 8 percentage points truly reflect the effect of a one-year increase in age at first marriage, we should observe two results. First, because a child born on March 31 enters elementary school at the same time as a child born on April 1, we should not observe a statistically significant discontinuity in age at first marriage at the April 1 cutoff, as there is no variation in school entrance timing. Second, we should not observe any discontinuity in the likelihood of childbirth at the April 1 cutoff, provided that we do not observe any variation in age at first marriage at this cutoff. Our falsification tests, therefore, examine whether the two arguments hold.

Figure 6 presents the graphical results of the falsification tests. The left three figures in Fig. 6 correspond to the density of age at first marriage by the running variable, and the right three figures correspond to the density of the share of females who have given birth by the running variable. The top two figures in Fig. 6 show the density of the two variables with a false cutoff equal to April 1. Thus, the figures compare those born in March with those born in April. As can be seen from the top figures, we do not observe any discernible discontinuity in age at first

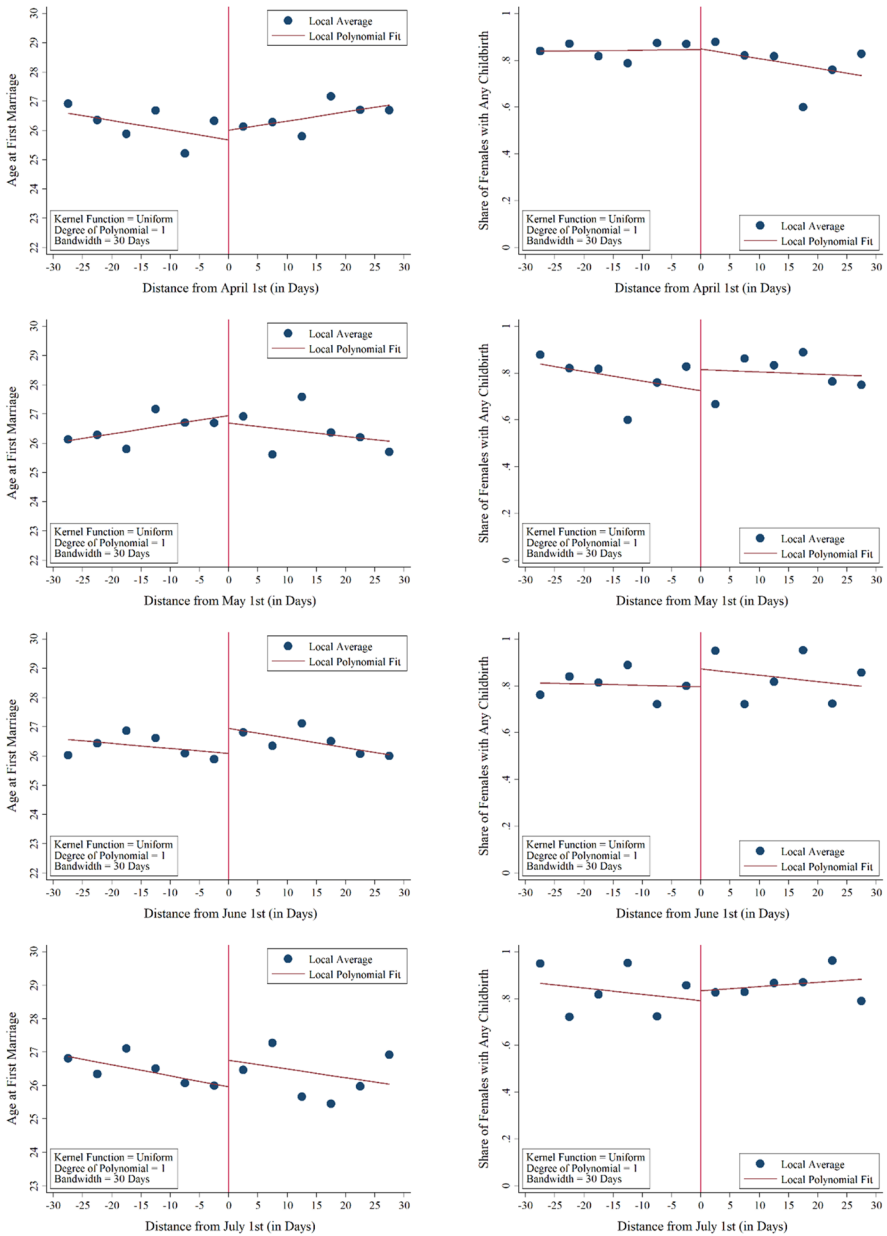


Fig. 6 Density of the treatment (left) and outcome variable (right) at false cutoffs

marriage at the false cutoff. We also do not observe any discontinuity in the share of females who have experienced childbirth. The subsequent figures in Fig. 6 show the density of these two variables at other false cutoffs (i.e., May 1, June 1, and July 1). While there are some variations in density by the false cutoffs, the graphical

Table 6 Results of Falsification Tests

Bandwidth choice	Age at first marriage		One or more births	
	Uniform	Triangle	Uniform	Triangle
Panel A: false cutoff at April 1st				
20 days	-0.552 (0.758)	-0.336 (0.839)	-0.106 (0.173)	0.457 (5.259)
30 days	0.141 (0.577)	-0.212 (0.654)	1.140 (60.734)	-0.032 (0.237)
40 days	0.008 (0.540)	-0.046 (0.565)	0.236 (1.364)	0.032 (0.389)
Panel B: false cutoff at May 1st				
20 days	-0.518 (0.949)	-0.179 (1.003)	0.279 (3.203)	-0.370 (0.525)
30 days	-0.329 (0.792)	-0.338 (0.849)	1.531 (38.210)	-0.351 (1.937)
40 days	-0.495 (0.652)	-0.364 (0.745)	-0.206 (0.385)	-0.217 (1.925)
Panel C: false cutoff at June 1st				
20 days	1.194* (0.673)	0.674 (0.716)	0.008 (0.066)	0.036 (0.118)
30 days	0.755 (0.581)	0.784 (0.565)	0.042 (0.103)	0.048 (0.095)
40 days	0.245 (0.441)	0.607 (0.471)	0.039 (0.163)	0.054 (0.111)
Panel D: false cutoff at July 1st				
20 days	0.532 (0.721)	1.293 (0.955)	-0.038 (0.049)	0.019 (0.093)
30 days	0.725 (0.569)	1.307* (0.673)	0.074 (0.152)	-0.011 (0.060)
40 days	0.353 (0.508)	0.864 (0.557)	0.499 (2.258)	0.032 (0.106)
Birth cohorts fixed effects	Yes			
Number of observations	1876			

Standard errors—clustered at the distance level—are in parentheses. Birth cohorts used in the analysis are those who are born from 1973 to 1982. The bandwidth choice of 20 days indicates the distance (in days) from March 1st. Age at first marriage—the treatment variable—in years. The outcome variable “One or More Births” is an indicator variable. The coefficient estimates for the outcome variable correspond to fuzzy regression discontinuity estimates with “Age at First Marriage” being the treatment variable. All the local polynomial regression estimators are conditional on baseline covariates presented in Tables 2 and 3

*Indicates statistical significance at the 10% level

results presented in Fig. 6 show, in general, that the magnitude of the discontinuity observed for the treatment and outcome variable at the false cutoffs is less salient than that observed at the true cutoff.

We examine the statistical significance of the estimated discontinuity at the false cutoffs in Table 6. Panel A shows the results obtained for the false cutoff at April 1. For the true cutoff, the estimated discontinuity in age at first marriage is around 1 to 1.5 years. For this false cutoff, the estimated discontinuity is significantly smaller (less than 0.5 years). Moreover, the estimated discontinuity varies to a great extent depending on the parameter choices, and none of the estimated discontinuity is statistically significant. Given that we do not observe any statistically significant discontinuity in the treatment variable, we should not observe any statistically significant discontinuity in an outcome variable. The estimates for the outcome variable are presented in the third and fourth rows of Table 6. As can be seen from the results, none of the estimates is statistically significant. In addition, the estimates again vary a great deal, indicating the noisiness of the data points. Finally, we obtain very similar patterns for the other false cutoffs. Of the 48 discontinuity estimates presented in the table, only two turn out to be statistically significant at the 10% level. Hence, we argue that the internal validity of our results is quite strong, given the results of the falsification tests.⁸

Conclusions

This study is, to the best of our knowledge, the first to provide a causal estimate of the effect of females' marriage age on fertility. We exploit Korea's school entry policy using exact date of birth and use an RDD to establish a causal relationship between age at first marriage and fertility. Specifically, we find that the school entry policy—which creates a difference of 12 months in school entrance dates between those born in February and March—delays females' age at first marriage by about 1 to 1.5 years. Using this plausibly exogenous variation in age at first marriage for the fuzzy RDD, we find that a year delay in age at first marriage reduces the probability of childbirth by about 8 to 10 percentage points. We also find that the same delay decreases the total number of childbirths by about 0.1 and 0.15 children. Moreover, we find evidence that one mechanism that drives such effect is the delay in age at first birth; we find that a year delay in age at first marriage leads to around 4 to 6 months of delay in age at first birth. Note, however, that further studies are desirable for the analysis of the total number of childbirths and age at first birth, because the effect estimates observed for these two outcomes are imprecise.

We believe that our study provides two policy implications. First, we argue that the very low level of fertility observed for many of the high-income countries in Asia and Europe is partly due to the increase in age at first marriage. While there are some variations, age at first marriage is increasing continuously in most of the countries facing

⁸ While not shown, the estimated discontinuity at other false cutoffs (e.g., August) is qualitatively similar to those shown in Table 6. These results are available upon request.

a very low fertility level. Average age at first marriage in countries with a high fertility level, on the other hand, has not increased significantly. The median age at first marriage for females in the United States, for example, was in the range of 25 to 27 between 2000 and 2010, while the TFR ranged from 1.93 to 2.06. While delay in age at first marriage reduces fertility, we argue that policymakers should not implement policies to prevent females from delaying marriage timing. Many studies have found that there are gains to delaying marriage (e.g., Loughran & Zissimopoulos, 2004). Hence, rather than engaging in advancing marriage timing, policymakers should engage more in identifying policy measures that allow the transition to adulthood and participation in higher-education institutions more compatible with motherhood and childbearing.

Second, we argue that it is very important to evaluate any policy as a collective (Lopoo & Raissian, 2012). Our study indicates that many policies that are not natalist policies may affect fertility, explicitly or implicitly, if such policies influence marriage age. For example, governments' financial aid policies aimed at promoting college completion rates are likely to delay marriage age. Similarly, many public policies may have unintended consequences with respect to fertility if they influence marriage age. Many policymakers argue that many of the natalist policies implemented worldwide are ineffective in raising the TFR. We argue, however, that the efficacy of such policies may have been diluted by the unintended consequences of many of the public policies that may have affected marriage age. Thus, we believe that increased effort should be expended on isolating the causal impact of existing natalist policies before making definitive conclusions regarding the efficacy of such policies.

Third, we refrain from drawing a definitive conclusion on the relationship between marriage timing and fertility from our findings. Because we used the Korean context to validate the statistical relationship between the two variables, the observed relationship may not be observed in other countries because Korea is unique in many aspects. For example, many of the potential factors such as gender roles that may mediate the relationship between fertility and marriage timing are different from other countries. Moreover, the instrument that we used in this study is school starting age. Because Korea is also different from other countries in terms of its education systems, the estimated effects of age at first marriage on fertility that are derived from such instrument may not be observed in other countries where education systems are quite different from Korea (i.e., local average treatment effects). For example, most females in Korea follow an educational sequence in a rigid manner. Therefore, the observed relationship between the two variables may not be salient in a country where an educational sequence is more flexible. All in all, we argue that one must be careful in generalizing the findings of this study to other countries that are quite different from Korea.

Finally, note that while the effect of age at first marriage on extensive margin is statistically significant, its effect on intensive margin is estimated imprecisely. We argue that more studies are necessary in particular as to its effect on intensive margin if we were to develop policies aimed at targeting the intensive margin.

Data availability The data used in this paper is not publicly available due to privacy issues. The authors will assist those who are interested in accessing the data.

References

- Balestra, S., Eugstger, B., & Liebert, H. (2020). Summer-born struggle: The effect of school starting age on health, education, and work. *Health Economics*, 29(5), 591–607.
- Bates, L. M., Maselko, J., & Schuler, S. R. (2007). Women's education and the timing of marriage and childbearing in the next generation: Evidence from rural Bangladesh. *Studies in Family Planning*, 38(2), 101–112.
- Becker, G. S. (1991). *A treatise on the family*. Harvard University Press.
- Black, S. E., Devereux, P. J., & Salvanes, K. G. (2011). Too young to leave the nest? The effects of school starting age. *Review of Economics and Statistics*, 93(2), 455–467.
- Dunson, D. B., Baird, D. D., & Colombo, B. (2004). Increased infertility with age in men and women. *Obstetrics and Gynecology*, 103(1), 51–56.
- Espenshade, T. J., Guzman, J. C., & Westoff, C. F. (2003). The surprising global variation in replacement fertility. *Population Research and Policy Review*, 22(5–6), 575–583.
- Feyrer, J., Sacerdote, B., & Stern, A. D. (2008). Will the stork return to Europe and Japan? Understanding fertility within developed nations. *Journal of Economic Perspectives*, 22(3), 3–22.
- Gelman, A., & Imbens, G. (2019). Why high-order polynomials should not be used in regression discontinuity designs. *Journal of Business and Economic Statistics*, 37(3), 447–456.
- Gutiérrez-Domènech, M. (2008). The impact of the labour market on the timing of marriage and births in Spain. *Journal of Population Economics*, 21(1), 83–110.
- Hahn, J., Todd, P., & van der Klaauw, W. (2001). Identification and estimation of treatment effects with a regression-discontinuity design. *Econometrica*, 69(1), 201–209.
- Higuchi, Y. (2001). Women's employment in Japan and the timing of marriage and childbirth. *Japanese Economic Review*, 52(2), 156–184.
- Ikamari, L. (2005). The effect of education on the timing of marriage in Kenya. *Demographic Research*, 12, 1–28.
- Imbens, G. W., & Lemieux, T. (2008). Regression discontinuity designs: A guide to practice. *Journal of Econometrics*, 142(2), 615–635.
- Isen, A., & Stevenson, B. (2008). Women's education and family behavior: Trends in marriage, divorce and fertility. *Topics in Demography and the Economy*. Cambridge: National Bureau of Economic Research Inc.
- Jones, G. (2007). Delayed marriage and very low fertility in Pacific Asia. *Population and Development Review*, 33(3), 453–478.
- Kabir, A., Jahan, G., & Jahan, R. (2001). Female age at marriage as a determinant of fertility. *Journal of Medical Sciences*, 1, 372–376.
- Kamb, R., & Tamm, M. (2022). The fertility effects of school entry decisions. *Applied Economics Letters forthcoming*. <https://doi.org/10.1080/13504851.2022.2039361>
- Lee, D. S., & Card, D. E. (2008). Regression discontinuity inference with specification error. *Journal of Econometrics*, 142(2), 655–674.
- Lee, D. S., & Lemieux, T. (2010). Regression discontinuity designs in Economics. *Journal of Economic Literature*, 48(2), 281–355.
- Lee, R., Mason, A. et al. (2014) Is low fertility really a problem? Population aging, dependency, and consumption. *Science* 346, 229–234.
- Lopoo, L. M., & Raissian, K. M. (2012). Natalist policies in the United States. *Journal of Policy Analysis and Management*, 31(4), 905–946.
- Loughran, D. S., & Zissimopoulos, J. (2004). Are there gains to delaying marriage? The effect of age at first marriage on career development and wages. Working Papers 207. RAND Corporation
- Marini, M. M. (1981). Effects of the timing of marriage and first birth on fertility. *Journal of Marriage and the Family*, 43(1), 27–46.
- McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics*, 142, 698–714.
- McCrary, J., & Royer, H. (2011). The Effect of female education on fertility and infant health: Evidence from school entry policies using exact date of birth. *American Economic Review*, 101(1), 158–195.
- Morgan, S. P. (2003). Is low fertility a twenty-first-century demographic crisis? *Demography*, 40(4), 589–603.
- Nahar, M. Z., Zahangir, M. S., & Shafiqul Islam, S. M. (2013). Age at first marriage and its relation to fertility in Bangladesh. *Chinese Journal of Population Resources and Environment*, 11(3), 227–235.

- Nguyen, H. T. M., & Lewis, B. (2020). Teenage marriage and motherhood in vietnam: the negative effects of starting school early. *Population Research and Policy Review*, 39(4), 739–762.
- Raj, A., Saggurti, N., Balaiah, D., & Silverman, J. (2009). Prevalence of child marriage and its effect on fertility and fertility-control outcomes of young women in India: A cross-sectional, observational study. *Lancet*, 373(9678), 1883–1889.
- Raymo, J. M., Park, H., Xie, Y., & Yeung, W. J. J. (2015). Marriage and family in East Asia: Continuity and change. *Annual Review of Sociology*, 41, 471–492.
- Seo, S. (2019). Low fertility trend in the Republic of Korea and the problems of its family and demographic policy implication. *Population and Economics*, 3(2), 29–35.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.