Is the association between education and fertility postponement causal? The role of family background factors

Abstract: A large body of literature shows a positive relationship between education and age at first birth. However, this relationship may in part be spurious due to family background factors that cannot be controlled for in most research designs. We investigate to what extent education is causally related to later age at first birth in a large sample of female twins from the UK (N=2,752). We present novel estimates using within-identical twin and biometric models. Our findings show that one year of additional schooling is associated with about half a year later age at first birth in OLS models. This reduced to only 1.5 months for the within-identical twin model that controls for all shared family background factors (genetic and family environmental). Biometric analyses reveal that it is mainly influences of the family environment – not genetic factors – that cause spurious associations between education and age at first birth. Lastly, we demonstrate using data from the Office for National Statistics that only 1.9 months of the 2.4 years of fertility postponement for birth cohorts 1944-1969 could be attributed to educational expansion based on these estimates. We conclude that (the rise in) educational attainment alone cannot explain differences in fertility timing (between cohorts).

---

9 This Chapter is based on an article by Tropf F. C. & J. J. Mandemakers (forthcoming). Is the association between education and fertility postponement causal? The role of family background factors, *Demography*
5.1 Introduction

A large body of literature from the USA and Europe documents a positive relationship between higher educational attainment and later age at first birth of women (e.g. Gustafsson, Kenjoh, & Wetzel, 2002; Lappegård & Rønsen, 2005; Martin, 2000). It has been argued that this relationship is causal and that it can account for the rise in the mean age at first birth during the educational expansion of the second half of the 20th century in Western countries (Balbo et al., 2013; Ní Bhrolcháin & Beaujouan, 2012). However, research also casts doubt on the idea of a causal effect of education on the age at first birth, suggesting that family background characteristics (social and/or genetic factors) cause spurious associations between educational attainment and fertility timing of women (Neiss et al., 2002; Rodgers et al., 2008). Social stratification research indicates that there is considerable similarity between parents-children and siblings in education and socio-economic attainment (Branigan, McCallum, & Freese, 2013; van Doorn et al., 2011) and in fertility behavior (Murphy, 1999; Rijken & Liefbroer, 2009). The similarity/intergenerational transmission may be due to parental socio-economic resources and/or socialization processes (Nisén & Myrskylä, 2014; Rijken & Liefbroer, 2009), but also to shared genetic dispositions, as there is sufficient evidence for a genetic component for both outcomes (Branigan et al., 2013; Mills & Tropf, 2016).

Our study contributes to the existing literature addressing the following three questions: first, does education indeed have a causal effect on age at first birth? Second, to what extent can the postponement of age at first birth during the second half of the twentieth century be explained by the simultaneous educational expansion? Third, to what extent are (social) environmental family background and/or genetic factors responsible for the observed relationship between education and age at first birth?

To answer these questions, we present within-identical twin models and engage in biometric modeling of the link between education and age at first birth in order to disentangle a possible causal relationship from genetic and environmental confounders in a large sample of twins from the UK. Next to quasi-experimental study designs, within-twin designs offer an approach to causality. Quasi-experimental designs (McCrary & Royer, 2011; Skirbekk, Kohler, & Prskawetz, 2006)
use exogenous factors, such as changes in compulsory school laws, which influence education but are supposed to be independent from fertility to investigate the causal effect of education on fertility. The within-twin approach, in contrast, controls for all factors shared amongst siblings, including genetic material, by using identical twins as a natural experiment (Kohler et al., 2011). The experiment is that the twins differ in levels of education, but neither in their genetic make-up nor their family environment. To the extent that the family background (social and/or genetic factors) is a common cause for education and age at first birth, standard regression models result in biased estimates, while within-(identical) twin estimates remain unbiased (Amin et al., 2015).

The within-identical twin approach is of particular interest for three reasons. First, as mentioned before, we expect that most of the unobserved factors important for education and fertility timing are related to the family of origin, which can be controlled for in twin studies (Branigan et al., 2013; Mills & Tropf, 2016). Second, the within-identical twin design can use variation across all levels of education, whereas for example quasi-experimental on school-law-change studies mostly rely on variation on a particular (mostly lower) level of education (see also Amin et al., 2015). Third, comparing fraternal and identical twins allows a quantification of the contribution of genetic and environmental influences for a specific trait, as in a classic-twin and behavior genetics modeling (Rodgers et al., 2007; Rodgers, Kohler, et al., 2001). Kohler and colleagues (2011) recently developed a new model, which formally integrates the two approaches: the within-identical twin regression model and the behavior genetics model. This so-called ACE-beta model represents a parsimonious solution to identify causality in a bivariate association and simultaneously evaluate the importance of genetics and shared environmental influences from the family.

Demographic research should strive to connect the individual and population level (Billari, 2015). We therefore combined our findings at the micro-level with nationally representative data from the Office for National Statistics (ONS) to evaluate whether the causal effect of education explained age at first birth trends at the population level. This was an important goal, not only in order to understand past fertility trends, but also to assess the potential to anticipate future fertility development based on changes in educational level - as education has gained
importance in fertility forecasting (Lutz, Butz, & Samir, 2014). Furthermore, in the UK, the trend in age at first birth was U-shaped during the mid-century in the UK (Hobcraft, 1996), whereas educational expansion increased steadily (Oreopoulos, 2006). These differential trends motivated a closer inspection of the relationship in different birth cohorts. We therefore also present analyses separately for birth cohorts born before and after the Second World War.

The current study builds upon and extends previous research in a number of ways. First, we applied the new ACE-beta model to study the link between education and age at first birth in a large sample of 2,752 twins from the TwinsUK. This is the largest adult twin register in the United Kingdom (Moayyeri et al., 2013) and it offers a larger sample size than was available to earlier research.

Second, previous investigations of the causal link between education and age at first birth in family designs showed mixed results, partly confirming a causal relationship (Amin & Behrman, 2014; Nisén et al., 2013) and partly not (Neiss et al., 2002; Rodgers et al., 2008). However, these studies also engaged in different analytical strategies, applying within-identical twin models (Amin & Behrman, 2014), standard behavior genetics models (Nisén et al., 2013) and complex structural equation models (SEM) (Neiss et al., 2002; Rodgers et al., 2008). To contribute to this puzzle we present within-identical twin models as well as the ACE-beta model that simultaneously estimates the causal effect and quantifies family environmental and genetic influences on the education-fertility relationship in a parsimonious SEM (Kohler et al., 2011).

Third, the UK is a particularly interesting case to study. Comparative studies within Europe suggest that the effect of education on age at first birth is relatively strong in the UK (Gustafsson et al., 2002; Rendall et al., 2005). A recent investigation by Ní Bhrolcháin & Beaujouan (2012) showed that in the UK between 1980-1999, 57 percent of the postponement of the mean age at first birth could be attributed to longer educational enrolment, while the remainder was due to additional, post-educational postponement effects – if we endorse the said causal relationship between education and age at first birth.

In the following, we first discuss the previous (bio-)demographic literature, continue by introducing the TwinsUK dataset, the within-twin method and the new ACE-beta model. We subsequently present our results and discuss our findings.
5.2 Background

5.2.1 Education and fertility timing

The mean age at first birth of women steeply increased by up to 4-5 years during the second half of the twentieth century all over Europe and the US, and was accompanied by an overall increase in educational attainment (Mills et al., 2011). In the UK, between 1980 and 2000, for example, the average age at first birth as well as age of leaving full-time education increased by 1.4 years (Ní Bhrolcháin & Beaujouan, 2012). Joshi (2002) showed by comparing three UK birth cohort studies (1946, 1958, 1970) that the percentage of women who were already mothers by age 26 dropped from 81 percent for those born in 1946 to 30 percent for the 1970 cohort. In all three birth cohorts, women with tertiary educational qualification were about half as likely to be mothers than women with no educational qualifications. In the most recent 1970 cohort, only 1 in 10 higher educated women were mothers at age 26, compared to 6 in 10 for women without qualifications.

A number of causal mechanisms have been put forward to explain the association between educational level and fertility timing (for an overview see Balbo et al., 2013; Mills et al., 2011). First, being enrolled in education itself may lead to postponement of childbearing, as it is difficult to combine the student role with the mother role, since both entail time intensive tasks. Women might delay childbearing due to high costs of children and few resources during the time of their studies or social norms might discourage parenting before the end of education (Hoem, 2000; Lappegård & Rønsen, 2005). Second, education may increase people’s aspirations and ability to pursue a career. So, women might further postpone childbearing until they are well established in their careers – also implying higher opportunity costs of the transition to parenthood given their larger human capital (e. g. Liefbroer & Corijn, 1999). Furthermore, education might also change values and orientations towards more individualistic lifestyles that seek fulfillment in life without children (Lesthaeghe, 1995; van de Kaa, 1987). Therefore, both educational enrolment and attainment may lead to postponing childbearing (and reducing fertility). Reverse
causality may play a role too because an early age at first birth may disrupt further education, although this is presumably mainly the case for unwanted teenage childbearing (Nisén & Myrskylä, 2014).

Reasoning along these lines and noting the often observed micro-level association led to the hypothesis that educational expansion in the UK explains the postponement of age at first birth (Ní Bhrolcháin & Beaujouan, 2012). However, such an interpretation is not straightforward for at least three reasons. First, the observed macro-level association between education and age at first birth changed across the past century; the simultaneous rise in age when leaving education and age at first birth only occurred after WWII. Second, in addition to the overall rise in education, there are a number of competing explanations for the rise in (female) age at first birth in developed countries. Third, as touched upon before, the relationship between education and age at first birth may be partly spurious. We now discuss these issues in more detail.

5.2.2 Changing macro trends

Comparative studies within Europe indicate that there is variability in the association between education and age at first birth (Gustafsson et al., 2002; Rendall et al., 2005). More specifically, in the UK the association between education and age at first birth on the population level changed across the past century. Age at first birth decreased in the period after the Second World War, and there was an accompanying steep increase in fertility - the so-called baby boom. At the same time educational attainment increased. The UK’s Education Acts from 1918 and 1944 lifted the school leaving age from 12 to 14 and later from age 14 to 15. As shown by Oreopoulos (2006), within three years during the period of 1945-1947 the fraction of 14-year-olds leaving schools fell from about 57 percent to less than 10 percent. Only later that century, since the 1960s, has age at first birth increased – the so called baby-bust. Education further expanded, and trends in age when leaving education and age at first birth ran parallel. Given the reversing trends throughout the century, we hypothesize that a strict causal logic that longer educational enrolment leads to later age at first birth may not apply.
5.2.3 Fertility postponement: Alternative explanations

Several alternative explanations - mainly related to the introduction of the ‘pill’ - have been put forward to explain fertility postponement in developed countries during the 20th century. It has been shown that rising female labor force participation (e.g. Rindfuss, Guilkey, Morgan, Kravdal, & Guzzo, 2007), ideational shifts in norms and values in sexual behavior and family planning of the second demographic transition (Lesthaeghe, 1995; van de Kaa, 1987) and increasing economic uncertainty (Andersson, 2000) are associated with an increasing age at first birth and might (partly) account for the postponement. The introduction of the ‘pill’ as an effective contraception is seen as an important trigger of these mechanisms.

In the UK, especially changes in family norms and values, as well as economic uncertainty, and the introduction of the pill provide alternative explanations for fertility postponement during the last 50 years, which may operate together with or independent of the educational expansion (Hobcraft, 1996; Murphy, 1993). To evaluate to what extent the educational expansion can be held responsible for fertility postponement and avoid ecological fallacy, we assess the causal effect of education on age at first birth on the individual level and use these estimates for predictions at the population level on nationally representative and independently collected data.

5.2.4 Critique of education and fertility link

The association between education and age at first birth might be in part spurious because unobservable factors influencing age at first birth may also be related to education. In particular, the family background may play a pivotal role as a socializing agent, a source of resources and support, and for transmitting genes. Studies have repeatedly demonstrated that education (Branigan et al., 2013) and fertility (Mills & Tropf, 2016) are associated with both the family environment and genetic factors. Therefore the family environment and genetic factors may (partly) explain the observed association between education and age at first birth.

Family background, for example, the socio-economic status of the parents, defines the resources and opportunities to remain in school longer and to financially
compensate for children. The status of parents can shape consumption and status aspirations of children who aim for higher education and social status in advance of family formation (Thornton, 1980). More generally, the socialized striving for autonomy might lower ambitions to build a family and increase investments in education and a career (Rijken & Liefbroer, 2009; Scott, 2004). Biological predispositions shared among family members and transmitted through genes can influence career and family trajectories. The timing of first attempts to get pregnant measured in retrospective interview is linked to one’s genetic make-up (Rodgers, Kohler, et al., 2001). It is also established that genes influence educational attainment (Branigan et al., 2013; Rietveld, Medland, et al., 2013), and fertility timing (Nisén et al., 2013; Tropf, Barban, et al., 2015). However, the question remains to what extent genetic effects for both outcomes are shared. In the following, we discuss previous studies that investigated the influence of (bio-)social family factors on education and fertility.

5.2.5 Biometric approaches for the education and age at first birth link

To our knowledge, four previous studies have considered both genetic and (socially) environmental family influences on the relationship between education and age at first birth. The first two investigations focus on the role of education as a possible mediator between cognitive ability and age at first birth. One used 813 pairs of relatives from the NLSY born between 1958-65 (Neiss et al., 2002) and the second 621 Danish twins pairs born between 1931-52 (Rodgers et al., 2008). They extend the classic twin model by introducing cognitive ability as preceding education in the causal chain at the cost of additional assumptions on shared genetic and environmental effects across traits. Both studies find that the observed mediating link of education turns non-significant after controlling for genetic and environmental influences from within the family. These two studies did not find genetic influences on age at first birth and thus no genetic correlation between both outcomes, which contrasts the general pattern (Mills & Tropf, 2016). They conclude that education may not directly delay childbearing, but individual differences which lead to higher cognitive ability and higher education inhibit fertility, and these differences arise between, not within families.
The third study estimated bivariate biometric models, which decomposed the (co)variance in education and age at first birth into latent genetic and environmental factors in a sample of 4,228 Finnish twins born in 1950-1957 (Nisén et al., 2013). For women they report that genetic, shared environmental factors of the twins as well as environmental factors unique to an individual independently explain part of the observed covariance between education and age at first birth. Influences unique to an individual can be interpreted as a causal effect (D’Onofrio, Lahey, Turkheimer, & Lichtenstein, 2013), but might also result from third factors influencing both outcomes at the same time such as the partner (Kohler & Rodgers, 2003).

Finally, Amin et al. (2014) used a within-identical twin design on a sample of 628 identical twins from the Minnesota Twin Registry. They find that one year of additional education leads to fertility postponement of around one year. They furthermore do not find evidence for family backgrounds effects on the association between education and age at first birth, that is, results in standard OLS regression models are nearly identical to the within-twin models.

In the current study, we first focus on the causal relationship between education and age at first birth and apply within-twin regression models. We present a within-identical twin estimate in a different context than Amin et al (2015), and for a larger sample. Second, we use our findings to project the age at first birth trend on the population level and contrast the projection with the actual age at first birth trend. Third, we extend these models according to Kohler et al. (2011) by integrating the bivariate twin model. Compared to previous mediation models (Neiss et al., 2002; Rodgers et al., 2008), the ACE-beta model needs fewer identifying assumptions and all paths we introduce to the model are identified and can be estimated.
5.3 Data & method

5.3.1 Data

We used information on twins of the TwinsUK registry. TwinsUK was originally established at the St. Thomas Hospital London in 1992 and gathered information on the life course of identical or monozygotic (MZ) and fraternal or dizygotic (DZ) twins (Moayyeri et al., 2013). Zygosity was established using standardized questions and confirmed by DNA genotyping in uncertain cases. Currently it contains information on about 12,000 individuals. We limited the analysis to same sex female individuals in complete twin pairs because the TwinsUK contains few male twins (<15 %) so that a comparable analysis for men was not feasible. We furthermore excluded women younger than 40 years old at last moment of observation to avoid an over-representation of young mothers and limit right-censoring of women who did not have children at the time of last observation. Valid information on zygosity, fertility and education was available for 3,856 women in 1,928 twin pairs. The sample is further reduced due to right-censoring of one or both twins in a twin pair to 2,752 women in 1,376 twin pairs. Note that for robustness checks, we also applied Cox regression models that include childless women (please see the SI 5-1, SI 5-2 and SI 5-3).

In order to compare the TwinsUK sample to a representative sample and to describe trends we used two additional sources from the ONS. For education, we used data from the General Household Survey (GHS) rounds from 2000 to 2006 (N=35,435, birth cohorts 1931-1970). The GHS is an annual continuous survey of the population in private households in Great Britain. To describe age at first birth we used estimates from the Office for National Statistics (Office National Statistics, 2013) because historical GHS fertility measures were partly limited to married individuals.

Age at first birth

The measure for age at first birth was based on information from two questionnaires in the TwinsUK. First, the ‘Main Questionnaire’, which was administered between 1995 and 2001, contains an inventory of the years of birth of
up to ten children. Second, age at first birth was assessed directly with the question, “How old were you when you had your first live birth?”, taken from the 2004 questionnaire. In case individuals participated in multiple waves, we used the earliest reported age.

Education

Education was measured as the age at leaving full-time education. This was assessed directly with the question: “At what age did you leave full-time education?” Most previous work has measured education in categories and imputed the corresponding years of schooling (e.g. Amin & Behrman, 2014; Rodgers et al., 2008). In the UK, it is difficult to order educational categories vertically due to differences in educational systems and migration across countries within the UK, especially because qualifications are so numerous, and because their names and content changes often (Jenkins & Sabates, 2007). Therefore, both TwinsUK and the GHS provide the age when leaving full-time education as a valid and reliable measure across data sources.

5.3.2 Analysis

The analysis proceeds in two steps. First, we present ordinary least square regression (OLS) and subsequently within-twin models (fixed effect models) to estimate the causal effect of education on age at first birth (AFB). Although within-twin models have a long history in economics (Rosenzweig & Wolpin, 1980) and psychology/behavioral genetics (D’Onofrio et al., 2013; Neale & Cardon, 1992), there are only some applications of within-twin models in demography (Amin & Behrman, 2014). A comparison of the OLS and the within-twin approach indicates to what extent family background leads to a spurious association between education and age at first birth. In a second step, we estimated the ACE-beta model to investigate to what extent genetic and/or environmental effects lead to a spurious association between education and age at first birth.
OLS and fixed effects models

The OLS-models include birth year and birth year squared of a twin (i) nested in a twin pair (j) to allow for a curvilinear trend in fertility timing across birth cohorts. Furthermore, we control for zygosity measured as the expected genetic relatedness between DZ twins (0.5) and MZ twins (1) and estimate robust standard errors to correct for the dependency structure of the twins (equation 1). We are interested in the effect of education ($\beta_1$).

$$afb_{ij} = \beta_0 + \beta_1 * edu_{ij} + \beta_2 * birth\ year_j + \beta_3 * birth\ year_j^2 + \beta_4 * zygosity_j + \varepsilon_{ij}$$

The OLS model gives a ‘naive’ estimation of the association between education and age at first birth. As detailed elsewhere (Amin, Behrman, & Spector, 2013; Kohler et al., 2011), standard OLS ignores at least three additional sources of variation; first the shared environment in a family j specific for fertility ($C_j$) such as family norms, second the individual additive genetic endowment ($A_{ij}$), and third the unmeasured unique environment ($E_{ij}$), which includes measurement error and individual specific effects for fertility, such as the influence of partner characteristics. These effects might impact the estimation of $\beta_1$ if they are correlated with education.

The within-twin models include fixed effects per family, which capture all (observed and unobserved) factors shared among the twins. We can discriminate between fraternal or dizygotic (DZ) and identical or monozygotic (MZ) twins. DZ twins share on average 50 percent of their segregating genetic material, MZ twins are genetically identical, so that a DZ-fixed effects model controls for half of all additive genetic effects (similar to a sibling fixed effects model) and the MZ-fixed effects model controls for all genetic effects. In other words, the MZ-fixed effects model controls for all factors shared amongst siblings, including birth year, zygosity, but also common (family) environment ($C_j$) and genetic effects so that $\Delta(C_{ij}) = 0$ and $\Delta(A_{ij}) = 0$. The equation for the MZ-fixed effects is:

$$\Delta(afb)_{ij}^{mz} = \beta_1 * \Delta(eedu)_{ij}^{mz} + \Delta\varepsilon_{ij}^{mz}$$

(2)
We applied both the OLS regression and the twin-fixed effects models to the pooled sample of twins, as well as for DZ and MZ twins separately. The comparison, particularly of the DZ and MZ models gives the first insights as to whether differences between the ‘naive’ OLS estimates and the fixed effects estimates are due to shared environmental and/or genetic factors. If genetic effects are important, we expect a stronger reduction in the effect of education for MZ-twins than for DZ-twins. The remaining link between education and age at first birth can be interpreted as causal. The effect is consistent if unique environmental influences important for education are independent of age at first birth (and vice versa) (Kohler et al., 2011).

Bivariate Genetic Modelling

Twin studies are no longer uncommon in social science (Branigan et al., 2013; Freese, 2008; Kohler et al., 2011, 1999; W B Miller et al., 2010; Mills & Tropf, 2016; Neiss et al., 2002; Nisén et al., 2013; Rodgers, Kohler, et al., 2001; Rodgers et al., 2008; Tropf, Barban, et al., 2015). Briefly: Twin studies are based on the comparison of MZ and DZ twins in order to quantify genetic and environmental influences. The degree to which MZ twins are more similar than DZ twins is assumed to reflect only genetic influences. Twin models are a typically SEMs that decompose the observed variance into three components: (a) additive genetic effects resulting from the sum of genetic effects from the whole genome (A), (c) environmental effects resulting from environmental influences shared between twins of a pair (C) and (e) non-shared environmental effects resulting from the unique environment of an individual (E) (which includes measurement error).

Three assumptions of the twin model need to be briefly addressed: The first is that MZ and DZ twins share their environment to the same extent (the equal-environment assumption (EEA)). This assumption has repeatedly been criticized (e.g., Horwitz et al. 2003), however, empirical evidence supports the validity of this assumption (Conley et al. 2013), including fertility studies (Felson, 2014). Second, it is assumed that there is no assortative mating within the population with respect to the outcome of interest. A violation of this assumption (Domingue and Fletcher 2014) would result in an underestimate of genetic influences. The third assumption
is that there are no non-additive genetic effects (gene-gene, and gene-environment interaction effects).

Following the same logic as in classic twin studies, it is possible to estimate the extent to which genetic and/or environmental factors are important for the covariance between two different outcomes. If education of twin 1 correlates with age at first birth of twin 2, then part of the covariance runs in families. If this correlation is higher amongst MZ-twin pairs than DZ-twin pairs, this indicates shared genetic effects for both outcomes. We estimate the ACE-beta model (Kohler et al., 2011), which represents an extension of a bivariate Cholesky decomposition (for details see Loehlin, 1996; Rodgers et al., 2007). The bivariate Cholesky decomposition estimates two equations, one for the first variable, in our case education (eq. 3), and one for the second variable, in our case age at first birth. Education of twin (i) is a function of:

$$edu_i = a_{edu,edu} * (A_{ij}^{edu}) + c_{edu,edu} * (c_{ij}^{edu}) + e_{edu,edu} * (E_{ij}^{edu})$$ (3)

Indexes indicate whether the effects are on education (edu), age at first birth (afb) or shared for both. Lower-indexes indicate whether the effects are common to both twins (j) or different (ij). The variance in education is therefore decomposed into additive genetic effects ($a_{edu,edu}$), shared environmental effects ($c_{edu,edu}$) and residual variance/measurement error ($e_{edu,edu}$).

Investigating the association between education and AFB, we formulate a second equation, which decomposes the variance in AFB into parts that are common with the genetic variance components of education (for example $a_{edu,afb} * (A_{ij}^{edu})$) and the shared environmental component ($c_{edu,afb} * (c_{ij}^{edu})$) and parts that are unique to AFB – caused by $A_{ij}^{afb}, C_{ij}^{afb}$ and $E_{ij}^{afb}$. In contrast to the classic bivariate Cholesky model, the ACE-beta model assumes that unique environmental sources of variance/measurement error for education ($E_{ij}^{edu}$) are uncorrelated with AFB ($e_{edu,afb} = 0$). Instead, the link between education and AFB is expressed by the causal link represented in a fixed-effects regression only within MZ-twin pairs $\beta_1 * \Delta(Ed)_{ij}^{mz}$ (eq. 4).
$$afb_{ij} = \beta_1 \cdot \Delta(e_{edu})_{ij}^{mz} + a_{edu,afb} \cdot (A_{ij}^{edu}) + c_{edu,afb} \cdot (C_{ij}^{edu}) +$$
$$+ a_{afb,afb} \cdot (A_{ij}^{afb}) + c_{afb,afb} \cdot (C_{ij}^{afb}) + e_{afb,afb} \cdot (E_{ij}^{afb})$$ (4)

In the main text, we visualize the standardized estimates in the form of a correlated factors model (Loehlin, 1996; Neale & Cardon, 1992) to facilitate interpretation. We present all results from these models and the standardization of the estimates in the SI. The genetic contribution to education and AFB is typically expressed as narrow-sense heritability ($h^2$), which is the proportion of variance due to additive genetic effects. In a parallel way we can compute the variance component for shared (C) and unique environmental influences/measurement error (E). Importantly, we give an estimate to what extent genetic and shared environmental influences important for education and age at first birth correlate and explain the observed correlation between education and age at first birth.

We used Stata 12 to estimate the OLS and fixed effect models and we graciously received the OpenMX R-function to estimate the ‘ACE-beta’ model from the developers Kohler et al. (2011).

Robustness

The main analysis presented in this paper included only twin pairs for whom both twins had a child because the bivariate biometric models cannot deal with non-linear outcomes. We tested the robustness of our results by including right-censored observations using Cox-regression models. We present Cox regression models with and without stratification by family to replicate the OLS and within-twin models for the full sample (Allison & Christakis, 2006). Please see the SI for details.

5.4 Results

5.4.1 Descriptive findings

Table 5-1 shows the descriptive statistics of the variables of interest separately for DZ and MZ twins. The twins in the sample were born on average just after WWII,
mean age at first birth was almost 26 and the mean age at leaving education was about 17 years for both kinds of twins. Most importantly for the biometric models, there were only minor differences in average or standard deviation of the outcomes of interest between DZ and MZ twins. It should also be noted that only 27 individuals (~1%) had their first birth before leaving education and only 74 (~2.7%) until one year after leaving education so that the temporal succession suggests that education influences age at first birth and not vice versa.

There are two main concerns about the within-twin approaches: first that variation of the independent variables within twin pairs is low and largely due to measurement error, and second that twin data are not representative. To address these concerns, first, Table 5-1 shows the mean absolute differences in education within twin pairs. These differences were substantial with 1.33 for DZ and 0.88 for MZ twins, which suggests that sufficient variation remained even within MZ twin pairs. The smaller differences for MZ than DZ twins suggests a genetic component underlying education. Second, TwinsUK is considered to be representative for the singleton population in the UK (Andrew et al., 2001; Moayyeri et al., 2013). Figure 5-1 shows the smoothed trends in age when leaving education and age at first birth for the representative ONS/GHS data (top panel) and the TwinsUK sample (lower panel). The general level and trends for the TwinsUK sample were similar to that of the ONS/GHS data, which increases our confidence in using the TwinsUK sample. However, it is to be noted that the mean age at first birth in the TwinsUK sample (25.83) was about a year earlier than in the ONS data (27.00). There was no difference in age at leaving education between the twinsUK (16.97) and GHS (16.92) estimates.

Moreover, it is clear from Figure 5-1 that the mean age of leaving education rose steadily throughout the 20th century, whereas the mean age at first birth followed a U-shaped pattern in the UK. Most previous research focused on trends in both outcomes during the second half of the 20th century, concluding that the rise in age at finishing education leads to a postponement of childbearing. However, the figure reveals a discontinuity in the association, which challenges the idea of a causal relationship between education and age at first birth.
Table 5-1. Summary statistics for monozygotic (MZ) and dizygotic (DZ) twins separately and within-twin absolute differences for education and age at first birth.

<table>
<thead>
<tr>
<th></th>
<th>MZ</th>
<th></th>
<th>MZ</th>
<th></th>
<th>DZ</th>
<th></th>
<th>DZ</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Min.</td>
<td>Max.</td>
<td>N</td>
<td>Mean</td>
<td>SD</td>
<td>Min.</td>
</tr>
<tr>
<td>Age at first birth</td>
<td>25.87</td>
<td>4.44</td>
<td>15</td>
<td>44</td>
<td>1354</td>
<td>25.79</td>
<td>4.66</td>
<td>16</td>
</tr>
<tr>
<td>Education (age leaving)</td>
<td>16.88</td>
<td>2.49</td>
<td>13</td>
<td>30</td>
<td>1354</td>
<td>16.90</td>
<td>2.67</td>
<td>12</td>
</tr>
<tr>
<td>Within-twin absolute differences</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age at first birth</td>
<td>3.41</td>
<td>3.27</td>
<td>0</td>
<td>22</td>
<td>677a</td>
<td>4.21</td>
<td>3.75</td>
<td>0</td>
</tr>
<tr>
<td>Education</td>
<td>0.88</td>
<td>1.80</td>
<td>0</td>
<td>15</td>
<td>677a</td>
<td>1.33</td>
<td>2.40</td>
<td>0</td>
</tr>
</tbody>
</table>

Notes:  

* refers to number of twin pairs. Education is measured as age when leaving full-time education.  

Source: TwinsUK, own calculations.
Figure 5-1 Trends in mean age at first birth and mean age when leaving full-time education in the TwinsUK sample and national representative data of the United Kingdom

Notes: AFB = age at first birth; The UK = estimates from large, nationally representative data from the Office for National Statistics
Source: TwinsUK and data from the National Office for Statistics (for details see methods section).

5.4.2 The causal effect of education on age at first birth

Our first research question asks: Is there a causal effect of education on age at first birth? Model 1 in Table 5-2 estimates the well-established naive OLS estimate of the effect of education on age at first birth at .44 (standard error (SE) = 0.04) in the present sample. This means that women who stayed one year longer in education had their first childbirth about half a year later on average. The effect holds independent of cohort effects as we controlled for birth year and its square. Model 3 and model 5 repeat the same model but now for subsets of MZ and DZ twins separately. The estimates of MZ and DZ twins were nearly identical.
Now we turn to the twin-fixed effects models to disentangle the causal effect from potential family confounders. Model 2 in Table 5-2 shows the pooled twin-fixed effects approach, which controls for all shared environmental factors among siblings and partly their genes. The effect of education on age at first birth fell to .11 (SE = 0.06), but remained significantly different from 0 at the 5 percent level in a one-sided test. Model 4 depicts the MZ-fixed effects estimate (0.12, SE = 0.09) that controlled for all factors that vary between twins including all shared environmental and genetic differences, as MZ twins are genetically identical. The twin-fixed effects estimate for DZ twins is depicted in model 6 (0.11, SE = 0.08). This estimate controlled for all factors that vary at the family level; genetic differences among DZ twins were only partly controlled for. The pooled, MZ-fixed effects, and DZ-fixed effects estimates were all similar in size (around .11), but the the two twin-fixed effects analyses by zygosity were not significantly different from 0. This was probably due to reduced power of the twin-fixed effects approach. This suggests a reduction in the causal effect of education on age at first birth with about three quarters (1-.12/.44) to about 1.4 months (.12*12 = 1.44). These findings suggest that a large part of the education effect on age at first birth can be attributed to family background factors (genetic and environmental).

In order to check the robustness of these findings when also including twin pairs where one or both women did not have children at the time of last observation and those younger than 40, SI 5-3 provides the results from the (stratified) Cox regression models (N = 4,398). Results followed the same pattern as the regression models except that the reduction in the estimated effect of education for the within-twin analyses appeared to be less strong. The probability of having a child decreased with each additional year of education by ~10 percent and this fell to ~3 to 4 percent in the stratified Cox regression models.
Table 5-2. Linear OLS and within-twin (FE) regression on age at first birth. Unstandardized estimates and standard errors in brackets.

<table>
<thead>
<tr>
<th>Model</th>
<th>Full sample</th>
<th>MZ twins only</th>
<th>DZ twins only</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OLS</td>
<td>Fixed effects</td>
<td>OLS</td>
</tr>
<tr>
<td>Education (age leaving)</td>
<td>0.44***</td>
<td>0.11*</td>
<td>0.46***</td>
</tr>
<tr>
<td></td>
<td>(0.04)</td>
<td>(0.06)</td>
<td>(0.06)</td>
</tr>
<tr>
<td>Year born</td>
<td>-3.79***</td>
<td></td>
<td>-5.19***</td>
</tr>
<tr>
<td></td>
<td>(0.78)</td>
<td></td>
<td>(1.13)</td>
</tr>
<tr>
<td>Year born squared</td>
<td>0.41***</td>
<td></td>
<td>0.56***</td>
</tr>
<tr>
<td></td>
<td>(0.08)</td>
<td></td>
<td>(0.12)</td>
</tr>
<tr>
<td>Zygosity</td>
<td>0.05</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.38)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>26.81***</td>
<td>23.98***</td>
<td>29.43***</td>
</tr>
<tr>
<td></td>
<td>(1.92)</td>
<td>(0.99)</td>
<td>(2.74)</td>
</tr>
<tr>
<td>N</td>
<td>2752</td>
<td>2752</td>
<td>1354</td>
</tr>
<tr>
<td>N pairs</td>
<td>1376</td>
<td>1376</td>
<td>677</td>
</tr>
</tbody>
</table>

Notes: * p<0.05, ** p<0.01, *** p<0.001, one-sided, OLS standard errors adjusted for non-independence of twins. Source: TwinsUK, own calculations.
Our second research question asked: to what extent can the educational expansion during the second half of the twentieth century explain the simultaneous overall postponement in age at first birth? Figure 5-2 depicts the simultaneous rise in age of leaving education (dashed line) and age at first birth (solid line) for cohorts born between 1944-1969. These cohorts started childbearing in time periods since the 1960s and therefore can be considered the main drivers of the fertility postponement. Age at first birth in 1969 in the UK (28.15) was about 2.4 years later than in 1944 (25.75). Figure 5-2 furthermore shows the simultaneous rise in age of leaving full-time education of around 1.3 years across these birth cohorts. A straightforward macro-explanation might consequently conclude that 54 percent (=1.3/2.4) of the 2.4 year of fertility postponement can be attributed to a rise in educational enrolment. However, as shown in the regression models, for each year of additional education, individuals postpone ~5.3 months in the OLS model and ~1.4 months in the within-twin models. Note that we use a reduced MZ-estimate of 1.2 (.10*12=1.20) months since we restrict the analysis to MZ-twins born after 1944 (not listed). Across birth cohort born between 1944-1969 only between 24 percent (=1.3*.44/24; OLS models) and 5.4 percent (=1.3*0.10/2.4) of the observed postponement in age at first birth can be directly related to the educational expansion. Figure 5-2 shows the predicted average age at first birth based on the OLS (short dash-dot-dot) and MZ twin-fixed effects results (dash-dot), as well as the explained (green) and unexplained (red) age at first birth trend across the second half of 20th century and in yellow the difference between OLS and MZ twin-fixed effects model. The OLS specification can be considered as the upper bound of the explained trend in AFB by education, and the MZ-fixed effects model the lower bound. to MZ-twins born after 1944 (not listed).

The estimates in Table 5-2 are based on birth cohorts born between 1919-1969. As shown in Figure 5-1, the correlation of the trends in education and age at first birth reversed on the population level and it is possible that the same applies to the individual level. Therefore, the causal effect of education might have become stronger in more recent cohorts. We additionally estimated models that allow for different educational effects for birth cohorts born before and after 1944 (the turning point in age at first birth trend) to investigate this issue (see SI 5-4). However, we do not find such differences.
Figure 5-2. Can educational expansion explain postponement of childbearing since 1945? The graph depicts the observed trends in mean age at first birth and education (age at leaving) for successive birth cohorts, and the predicted trend in age at first birth by education using the OLS and MZ estimates.

Notes: The trend for educational level was rescaled to that of age at first birth. Predictions were based on estimates from Model 2, Table 5-2, and data from Table 5-1.

Source: General trends in education were derived from British General Household Surveys (2000-2006) and age at first birth from the Office for National Statistics, Cohort fertility, Table 2. (Office National Statistics, 2013)
5.4.3 The role of genes and the environment as family background factors

The reduction of the estimate of the causal effect of education on age at first birth in OLS versus MZ twin-fixed effects models leaves us with the third research question: To what extent is this due to unobserved shared environmental influences and/or unobserved genetic endowments that are common to both outcomes. In order to answer this question, we present the ACE-beta model, which simultaneously estimates the direct causal link between education and age at first birth based on the MZ twin-fixed effects estimator and the genetic and shared environmental contribution. Figure 5-3 presents the correlated factor model (please see SI 5-5 and 5-6 for the untransformed SEM estimates and the transformation).

The estimates are well in line with previous studies (Branigan et al., 2013; Mills & Tropf, 2016). The heritability of education was 0.46, meaning that 46 percent of the variance in education was associated with additive genetic differences. The shared environmental influences account also for 23 percent of the variance, and the unique environment, which includes measurement error, for 31 percent. The heritability of age at first birth is estimated to explain 32 percent of the variance in age at first birth. Shared environmental effects were smaller (0.12) but significant and unique environmental effects/measurement error, made up the largest part (0.56).

All variance components were statistically significant. We also see the correlation of genetic and environmental influences across outcomes. If the genetic/environmental correlation between two traits is 1, all genetic/environmental variance in trait 1 and 2 has a common base. If the correlation is 0, the genetically/environmentally based variance between trait 1 and 2 are independent. Shared environmental effects on education and age at first birth correlated to 1 meaning that there are no shared environmental effects for age at first birth independent of shared environmental influences on education. Also a model that constrained shared environmental effects for education and age at first birth to be the same had an equal statistical fit as the model freely estimating the parameter (not listed). Importantly, genetic effects across both outcomes do not significantly correlate (0.14), implying that family background factors causing a spurious association between education and age at first birth were mainly environmental in
nature. Note that the direct effects from the MZ twin-fixed effects estimates reduces to 0.07 in this model specification. In the ACE-beta model, the direct link between education and age at first birth explained 16 percent of the association, shared environmental influences amongst siblings 62 percent, and shared genetic influences 22 percent (see Si 5-5 and SI 5-6 for details).

Figure 5-3 Correlated factor model depicting standardized estimate of genetic (A) and shared environmental (C) influences on education and age at first birth, the correlation of these influences across traits as well as a causal effect of education on the age at first birth.

Notes: * p<0.05, ** p<0.01, *** p<0.001, two-sided
Source: TwinsUK, own calculations.
5.5 Discussion

The supposed causal effect of education on fertility postponement has become a parsimonious and powerful explanation for demographers, not only for the fact that higher educated women have their first baby at later ages, but also for the general postponement of childbirth during the second half of the 20th century due to the educational expansion (Ní Bhrolcháin & Beaujouan, 2012). In this study, we challenge the claim that education causally influences age at first birth and consequently that the educational expansion is the main cause of recent fertility postponement. We present within-twin (fixed effects) and novel biometric models using a unique dataset of female twins from the UK. We estimate the causal effect of education on age at first birth, and the extent to which environmental and genetic factors cause a spurious association. We find a reduced effect of education on age at first birth in the within-twin design compared to standard regression models. Complementary biometric analyses reveal that the association between education and age at first birth is to a large extent caused by social family background effects, whereas genetic inheritance plays only a small role. These results suggest that the prevailing view of a strong causal effect of education has no merit and needs to be revised.

In our study, the effect of all family factors shared among siblings explained up to two thirds of the observed association between education and age at first birth, and only a small part of the effect is due to genetic dispositions. This finding provides a strong justification for continued research into the role of family background effects on education and fertility. However, the twin models we present do not give insights into which family background factors are important for both education and age at first birth. Previous investigations have pointed to socio-economic status of the parents (van Doorn et al., 2011) and parental demographic behavior (e.g. Lappegård & Rønsen, 2005; Marini, 1985; Nisén & Myrskylä, 2014; Rijken & Liefbroer, 2009). Recent investigations, also show that social ties such as friends (Balbo & Barban, 2014) influence fertility timing and that siblings influence each other (Lyngstad & Prskawetz, 2010). Given that measured family characteristics typically only account for a small part of the explained variance (e.g. Nisén & Myrskylä, 2014), and parents are important in choosing who to befriend and date,
one direction to better understand the role of the family in fertility behavior is in further study of social and network dynamics (of the family).

Since the central goals of demographic research are the discovery, description and explanation of phenomena at the population-level, individual-level approaches have been criticised for their inattention to the core perspective in demography (Bhrolcháin & Dyson, 2007; R. Lee, 2001; Smith, 2003). Ní Bhrolcháin and Dyson (2007) proposed a catalogue of ten criteria for making causal claims when linking macro phenomena such as the educational expansion to the postponement of fertility. These include: the temporal order of events, the no-cause-no-effect-rule, the proposition of mechanisms, and the proportionality of the scale of cause and effects. In a comparable study with nearly identical findings at the population level to those of our study, Ní Bhrolcháin and Beaujouan (2012) assign an important role to education for fertility postponement, which underscores the importance of education for fertility projections (Lutz et al., 2014).

A central goal of our study was to link our findings at the individual level to the population level (Billari, 2015). The main ways in which our approach is distinct are that the unit of analysis changes from populations to individuals and that we present and ‘exception’ of the no-cause-no-effect rule, since we postulate that fertility postponement is not a mono- but a multi-causal phenomenon. Our findings mainly challenge those from Ní Bhrolcháin and Beaujouan (2012) with respect to the criterion of proportionality, as we find that increasing educational attainment can only explain a fraction of the trend in age at first birth—if we integrate results from the individual level into our projection. Perhaps the educational expansion serves as a proxy in many studies for other simultaneous historical developments that are not directly measured. An extensive review by Hobcraft (1996) identified changes in sexual norms and family planning, the dramatic economic crises during the 1970s and 80s, and the introduction of the ‘pill’ as reasons for declining fertility levels and fertility postponement within the UK. In that sense, we agree that, for the relationship under study, the question is not ‘what causes fertility postponement?’, but ‘to what extent does the educational expansion (and other factors) cause fertility postponement?’. Overall, the role of education as a useful predictor of future fertility trends is in doubt.
Previous investigations on both the causal effect of education on age at first birth and the role of genetic versus environmental family background factors produced mixed findings, and our study may contribute to explaining these inconsistencies. First, the nuanced interpretation of the role of education fits with two previous studies. Neiss et al. (2002) and Rodgers et al. (2008) show for the US and Denmark that the association between education and age at first birth is completely absorbed by latent family influences. Nisén and Myrskylä (2013) and Amin and Behrman (2014), however, find a significant link between both outcomes net of family influences. Our study suggests that a causal effect of education on age at first birth exists, however, that it is small, and therefore might have remained undetected in former investigations. Note that in the current study, the causal effect estimated from within-identical twins – and therefore also in the ACE-beta model – is not statistically distinguishable from zero, and only if we pool identical and fraternal twins – which have nearly identical estimates – do we detect the causal link.

It also needs to be mentioned that findings on the role of the genetic component in age at first birth in general are mixed in the literature. We find that around one third of the variance in age at first birth is explained by additive genetic effects. This is similar to e.g. studies from the US (Byars et al., 2010), Australia (Kirk et al., 2001) or Finland (Nisén et al., 2013). However, other investigations from the US (Neiss et al., 2002) and Denmark (Rodgers et al., 2008) find no significant genetic influences. Differences in genetic effects on fertility across countries and within countries over time exist – also within the UK (Tropf, Barban, et al., 2015) - and may be due to gene-environment interaction (Kohler et al., 2006; Mills & Tropf, 2016). Further investigation into the genetic pathways to fertility is needed, also in order to better understand patterns of gene-environment interaction across populations – also including molecular genetic information (Tropf, Stulp, et al., 2015).

The within (identical) twin approach we applied forms a useful tool to establish a causal relationship between two variables yet it has its limitations. The most critical assumption is presumably that variation in education within twin pairs is uncorrelated with variation in age at first birth. The question arises: what makes (identical) twins different in educational attainment? If for example health issues lead to early school dropout and delayed fertility, the causal effect estimated from
the within twin models will be smaller than the true effect. If the presence of a partner influences career and fertility aspirations towards longer educational and later fertility, the true causal effect would be smaller than the estimate from the within-twin models. Additionally, measurement error can downwardly bias the estimates (see for discussion Amin et al., 2013). However, conservative predictions based on estimates from the OLS models suggest that the educational expansion cannot account for a large part of observed fertility postponement. Nevertheless, we advise caution for the interpretation of our findings insofar as we find mean differences in age at first birth across TwinsUK and the GHS data. Alternative designs can validate our findings in future research. Desirable approaches include quasi-experimental designs using instrument variables (for a recent discussion see Amin et al., 2015; Boardman & Fletcher, 2015).

In general, our study challenges the common approach of explaining differences and trends in fertility timing mainly by invoking educational differences. Differences between families and societal changes and upheavals across time can have a strong impact on fertility timing, which may be mistakenly attributed to the causal influence of just one factor. This study shows that fertility timing is the result of a complex interplay of environmental and genetic influences. We conclude that (the rise in) educational attainment alone cannot explain differences in fertility timing (between cohorts).
Literature


Ioannidis, J. P., Ntzani, E. E., Trikalinos, T. A., & Contopoulos-Ioannidis,


methods for twins data: Integrating causality, endowments, and heritability. *Biodemography and Social Biology, 57*(1), 88–141.


Lee, S. H., Yang, J., Goddard, M. E., Visscher, P. M., & Wray, N. R.


Rietveld, C. A., Cesarini, D., Benjamin, D. J., Koellinger, P. D., Neve, J.-E.


Environment, 3(1), 51–72.


Visscher, P. M., Brown, M. A., McCarthy, M. I., & Yang, J. (2012). Five


