Routledge
Taylor & Francis Group

Population Studies
A Journal of Demography

ISSN: 0032-4728 (Print) 1477-4747 (Online) Journal homepage: http://www.tandfonline.com/loi/rpst20

Genetic influence on age at first birth of female twins born in the UK, 1919–68

Felix C. Tropf, Nicola Barban, Melinda C. Mills, Harold Snieder & Jornt J. Mandemakers

To cite this article: Felix C. Tropf, Nicola Barban, Melinda C. Mills, Harold Snieder & Jornt J. Mandemakers (2015) Genetic influence on age at first birth of female twins born in the UK, 1919–68, Population Studies, 69:2, 129-145, DOI: 10.1080/00324728.2015.1056823

To link to this article: http://dx.doi.org/10.1080/00324728.2015.1056823

Population Studies, 2015
Vol. 69, No. 2, 129–145, http://dx.doi.org/10.1080/00324728.2015.1056823

Genetic influence on age at first birth of female twins born in the UK, 1919–68

Felix C. Tropf[1], Nicola Barban[1,2], Melinda C. Mills[1,2], Harold Snieder[1] and Jornt J. Mandemakers[1,3]

[1]University of Groningen; [2]University of Oxford; [3]University of Wageningen

Using a sample of monozygotic (945, 42 per cent) and dizygotic (1,329, 58 per cent) twin pairs born 1919–68 in the UK, we applied innovative tobit models to investigate genetic and environmental influences on age at first birth (AFB). We found that a substantial part (40 per cent) of the variation in AFB is caused by latent family characteristics. Genetic dispositions (26 per cent) play a more important role than the shared environment of siblings (14 per cent), with the non-shared environment/measurement error having the strongest influence (60 per cent). Like previous studies, this study reveals marked changes in estimates over time, and supports the idea that environmental constraints (war or economic crisis) suppress and normative freedom (sexual revolution) promotes the activation of genetic predispositions that affect fertility. We show that the exclusion of censored information (i.e., on the childless) by previous studies biased their results.

Supplementary material for this article is available at: http://dx.doi.org/10.1080/00324728.2015.1056823

Keywords: age at first birth; fertility; UK; heritability; heredity; reproductive age; gene–environment interaction; twins

[Submitted January 2014; Final version accepted March 2015]

Introduction

There has been a steep increase in age at first birth (AFB) across Europe since 1970. It is now between 28 and 29 years of age (Mills et al. 2011), and is often described as the 'postponement transition' (Kohler et al. 2002a). The primary reasons proposed for postponement have been women's increased inclination to obtain higher education and employment in the labour force, but cultural transformations surrounding the timing and role of children have also been proposed as a reason (Van de Kaa 1987; Balbo et al. 2013). Attempts to explain differences and changes in AFB patterns within demography and the social sciences have continued to rely almost exclusively on associations with socio-environmental factors (Hobcraft 2006; Goldstein et al. 2009; Mills and Tropf 2015). Yet a growing body of literature has indicated an interplay between individual genetic endowment and the environment, which in turn shapes subsequent life-course decisions and fertility behaviour (Udry 1996; Murphy 1999; Guo 2005; Kohler et al. 2006; Freese 2008; Mills and Tropf 2015). Significant genetic influences related to fertility have been established for age at first dating or marriage (Mealey and Segal 1993), age at first sexual intercourse (Dunne et al. 1997), number of sexual partners (Guo et al. 2008), and the likelihood of unprotected sexual intercourse (Daw and Guo 2011; see Harden 2013 for a review of genetic influences on sexual behaviour). One of the strongest genetic effects was found in a twin study of women's (self-reported) first attempt to have a child. The study found that 40–50 per cent of the variance was explained by genetic differences (Kohler et al. 1999; Rodgers et al. 2001).

A prominent technique used to determine the relative influence of genes and environment in explaining AFB is a twin study. Twin data allow comparisons between monozygotic (MZ), that is, identical, and dizygotic (DZ), twins and are therefore one of the best resources for evaluating the importance of genetic variation in observed traits (Boomsma et al. 2002). However, existing research that adopted this design to examine AFB has produced very mixed findings. Whereas an Australian study found

significant genetic influences (Kirk et al. 2001), research in the US (Neiss et al. 2002) and Denmark (Rodgers et al. 2008) reported non-significant effects (Mills and Tropf 2015). The aim of the study reported here was to extend existing knowledge about the influence of genetic and environmental factors on AFB in two ways: by replicating previous research on a new population of twins in the UK, and by introducing new analytical techniques that can explain different findings by including or excluding childless individuals. The study addressed three central questions. What are the relative effects of genes and environment in explaining differences in AFB in the UK? Did the pattern of explanation change for cohorts born over the twentieth century? How does the modelling procedure used affect the results obtained?

The current study built upon but also extended previous research in three main ways. First, we introduced more sophisticated statistical models and demonstrated that they produced results different from those of classic designs. The core difference was in the handling of right-censored observations (i.e., on individuals who had not had a first birth when last observed). Reviving awareness of the censoring issue in behavioural genetics (Kohler and Rodgers 1999), we examined whether the inclusion of censored information affected heritability estimates of the AFB and if so, to what extent. This was achieved by introducing a tobit structural equation model (SEM) into the behavioural genetic framework, applying it with censored cases included, and then comparing the findings with estimates obtained using the classic linear SEMs—standard in behavioural genetics—that excluded censored cases.

A second contribution was the application of local linear regressions for each birth cohort in order to describe underlying trends more explicitly (Cleveland 1979). This allowed us to explore changes over the past century in the influence of latent factors contributing to AFB. Kohler et al. (1999, 2002b) having demonstrated that the heritability of number of offspring depended strongly on historical and demographic changes during the past century. Our study used the rich information from a large twins register, TwinsUK, in the UK, which contains information on birth cohorts from 1919 to 1968. It was essential to examine cohort differences because the twentieth-century UK was the setting for distinct social upheavals known to have a strong impact on AFB (Murphy 1993; Hobcraft 1996, 2006).

A third and related substantive contribution was the study of this topic with a new population of UK women. Human traits and behaviour are the result of interplay between genetic and environmental factors (Freese 2008). Genes provide predispositions for complex traits such as AFB, but environmental conditions determine whether such dispositions will be activated. To properly understand gene–environment interactions across societies and time, or in other words the context dependency of genetic effects on fertility, we must therefore evaluate the estimated heritability of a trait in different environments (Guo 2005). The UK is an interesting case since intergenerational correlations of fertility and AFB are well established and family background has been shown to be a key predictor of fertility over the twentieth century (Booth and Kee 2009). It was therefore important to understand the relative influence of genetic and environmental factors on this association.

TwinsUK, the source of our data, is the largest adult twin registry in the UK. It contains data on 12,000 twins and is primarily used to study the genetic and environmental aetiology of age-related complex traits and diseases (Moayyeri et al. 2013). We examined women only, using data available for 2,274 female twin pairs (and thus 4,548 individuals). In the following section we briefly introduce behavioural genetics concepts. We then review previous research on familial and genetic effects on fertility and the course of fertility in the UK over the twentieth century. We continue by introducing the TwinsUK data set and our methods, and follow this with a presentation and discussion of the central findings.

Background

Intergenerational correlations of fertility

The motivation to disentangle effects of shared genetic factors from shared family environment on fertility has early roots (for a review, see Mills and Tropf 2015). Fisher (1930) conducted a study on the number of offspring of more than 2,000 British aristocrats at the end of the nineteenth century and found an intergenerational correlation of 0.20. Murphy (1999) concluded that intergenerational correlations of fertility level were very common and had even increased since the end of the nineteenth century.

Although intergenerational correlations of fertility level seem to follow a consistent pattern, their interpretation remains controversial. One explanation might be that parents each transmit 50 per cent of their genetic variants and thus genetic

predispositions to their children. With this in mind, Fisher (1930) interpreted the intergenerational correlation of 0.20 as exclusively a genetic effect, which meant that 40 per cent of the variance was explained by genetics. Others (Booth and Kee 2009) argued that parents also transmit environmental characteristics such as socio-economic status to their children, which are also important for fertility behaviour. Thus, intergenerational correlations between the fertility levels of parents and children lack established interpretations of their genetic and environmental origins, and alternative designs such as twin studies are required to quantify their effects.

Genetic research on human fertility

The majority of fertility research that has employed a twin-study design has used data from a historical Danish Twin registry that includes information on virtually every pair of twins born since 1870. Using these data, Kohler et al. (1999, 2002b) found evidence of a strong genetic influence on number of offspring. They report that the influence was particularly strong following the First and Second Demographic Transitions. The heritability of the number (genetic variance in the number of offspring as a proportion of population variance in the number) was a modest 0.40 during the strong fertility decline of the First Demographic Transition at the end of the nineteenth century and the Second Demographic Transition in the second half of the twentieth century, but was close to zero between the two transitions.

It is the large changes in heritability after the demographic transitions that support the hypothesis of gene–environment interaction. Kohler et al. (1999, 2002b) associate these changes in genetic influence with the environmental changes that occurred during the transitions. In particular, there were improvements in economic, medical, and hygiene conditions during the First Demographic Transition, while the Second Demographic Transition saw the introduction of oral contraceptives that enabled, and cultural transformations that encouraged, the exercise of personal choice over childbearing (Lesthaeghe and Van de Kaa 1986; Van de Kaa 1987; Lesthaeghe 1995). According to the standard rationale, the more childbearing (whether and when) is subject to personal choice in a society, the more it is influenced by genetic predispositions (Udry 1996). The introduction of the contraceptive pill during the Second Demographic Transition is believed to have had a particularly important effect by offering new freedom over the timing of the first child (Van de Kaa 1987; Lesthaeghe 1995).

The same pattern of gene–environment interaction is suggested by a study of the Danish sample that examined the number of offspring of members of birth cohorts at the age at which 25 per cent of them had a first child (Kohler et al. 2002b). This measure of the level of early fertility showed a heritability of 0.52 for cohorts born in the period 1961–68 and therefore socialized after the Second Demographic Transition, when norms encouraging early childbearing had weakened. For the cohorts born in the period 1945–52, this effect was close to zero.

Unlike the fertility studies, investigations of genetic and environmental influences on AFB have shown mixed results. Neiss et al. (2002) conducted a study in the US using constructed kinship data from the National Longitudinal Survey of Youth (NLSY). Cohorts born in the period 1958–65 were found to have a low heritability of 0.06 for AFB, whereas the shared environment explained up to 20 per cent of the observed variance. Similar results were found in Rodgers et al.' s (2008) study of twins from the Danish sample. For birth cohorts for the period 1931–52 (1,242 twins), the shared environment explained 26 per cent of the variance in AFB and no genetic effect was found.

In contrast, an Australian twin study reported a significant heritability of AFB. Kirk et al. (2001) investigated the extent of genetic influence common to age at menarche, age at first birth, age at menopause, and number of offspring (N = 2,710 twins). The results showed that for cohorts born in the period 1900–65, 21 per cent of the observed variance was explained by genetic variance. The shared environment explained 18 per cent of the observed variance, and the non-shared environmental variance/measurement error accounted for 61 per cent. At the same time, their multivariate biometric models showed that genes influencing AFB were associated with overall reproductive success—measured as the number of children of an individual in relation to the average number of children in a birth cohort—and therefore of the survival of the genetic variants.

A recent investigation from Finland (Nisén et al. 2013) of cohorts born in the period 1950–57 shows similar results. Genes explained 26 per cent of the variance in AFB, shared environment 12 per cent, and the non-shared environment/measurement error 61 per cent. There was also a genetic correlation between education and AFB. In other words, the

genetic effects on both outcomes had a common base that partly explains the observed covariation.

Overall, the studies suggest that the influence of genes and the environment on AFB depends on time and place. The treatment of censored information in the estimation of these variables might also play a crucial role. While the studies from the US, Denmark, and Finland included only individuals who experienced a first birth, the Australian models also included childless individuals by imputing the age at last observation, grouping the ages, and estimating an ordered SEM. For this reason, we applied both standard models, which exclude childless individuals, and models that have the ability to capture censored cases. We did this so that we could compare results for the same population.

In what follows, we first briefly describe the fertility context of the UK during the twentieth century, and then the TwinsUK register we used. In contrast to previous historical research on genetic effects on reproductive behaviour, our study focused entirely on AFB. However, historically AFB has been strongly related to number of offspring, and changes in this number during the twentieth century have been accompanied by postponement of childbearing (Mills et al. 2011; Balbo et al. 2013). We therefore expected our study to reveal that AFB was subject to patterns of genetic and environmental influences similar to those found for the number of offspring.

Fertility context of the UK in the twentieth century

Historical analysis on the genetics of fertility focused on the analysis of birth *cohorts* to reduce environmental heterogeneity among individuals. To interpret the dependence of genetic and environmental contributions to AFB on period, however, we had to examine the historical *periods* in which the cohorts started childbearing. Figure 1 shows the trend in number of offspring and AFB for birth cohorts in the UK between 1925 and 1970. We observe the well-established trends: first an increase in the average number of offspring for birth cohorts born during the 1920s and 1930s, followed by a steady decline in completed fertility during the second half of the past century. The AFB shows in general the inverse pattern: a decrease in the average AFB for birth cohorts across the 1920s and 1930s, a fall at the beginning of the 1940s, and the striking postponement in the AFB for cohorts born after the end of the Second World War. These trends have been

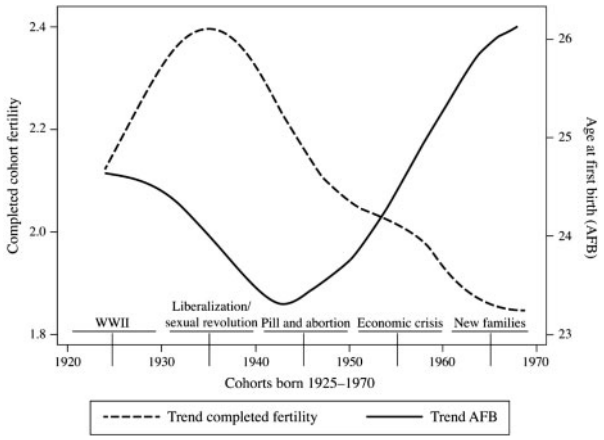


Figure 1 Smoothed scatterplot of completed cohort fertility and AFB in England and Wales, cohorts born 1925–70

Sources: AFB UK: the estimated average AFB in the UK. Because, historically, official data on birth order have been only collected within marriage, these values are based on estimates from the Office for National Statistics, Cohort fertility, Table 2. Available at: http://www.ons.gov.uk/ons/publications/re-reference-tables.html?edition=tcm%3A77–263133 (accessed March 2013). Completed cohort fertility UK: Human Fertility Database, Max Planck Institute for Demographic Research (Germany) and Vienna Institute of Demography (Austria). Available at: www.humanfertility.org (accessed 1 December 2013).

analysed and their direction related to at least five historical periods of social change and upheaval in the UK that have been strongly related to fertility behaviour (see Figure 1 for illustration, and Murphy 1993; Hobcraft 1996, 2006).

In the first of the five periods we considered, British women born in the 1920s grew up and reached maturity during the Second World War, a time of major environmental constraints that affected the timing of AFB: food rationing that began in January 1940; the 'Battle of Britain' in September of the same year; and the call in March 1941 for 100,000 women (but not married women with young children) to 'come into the factories'. Hobcraft (1996) argues that their circumstances made women in these birth cohorts reluctant to start childbearing and they postponed doing so.

In the second period, women born in the 1930s benefited from the new welfare provisions introduced after the war, such as a free National Health Service that included maternity care, family allowances, and free secondary education to 'feed the aspirations of the middle classes' (Hobcraft 1996,

p. 495). Furthermore, the *period* of the 1950s—when these birth cohorts started childbearing—was a period of changing mores, culminating in the sexual revolution of the 1960s (Hobcraft 1996). There was a growing insistence on the right of individuals to make their own decisions about sexual conduct and family formation, though the use of reliable contraception was only spreading slowly.

Birth cohorts born in the 1940s and 1950s, the subjects of our third and fourth periods, started childbearing during the 'baby bust' of the 1960s and 1970s, one of the most critical periods for understanding contemporary fertility developments. Two crucial factors influenced fertility in these two periods. One was the introduction of effective contraception together with the legalization of abortion laws. According to Hobcraft (1996), more than half the fertility decline in the period can probably be attributed to the reduction in the number of unwanted children. The other important influence on the third and fourth periods was the economic crisis of the 1970s. At the end of 1972, wage and rent freezes dominated the UK, and their effect was reinforced by the rise in oil prices in 1973. House prices rose rapidly in 1974 and unemployment passed the 1 million threshold in 1976, growing to 3 million in 1982 (Hobcraft 1996). The proportion of married women who had ever used the pill during the first 5 years of marriage increased from less than one-third in 1967 to around 80 per cent in 1976 (Murphy 1993). People were able to control childbearing and effectively respond to this period of economic insecurity by postponing childbirth. In fact, England and Wales reached a historical low point in total fertility in 1977 (Hobcraft 1996).

The subjects of the fifth and final period are the birth cohorts born in the 1960s, the members of which entered their 20s during the partnership revolution of the 1980s (Kiernan 2004). Cohabitation became more socially accepted as a form of partnership and as one able to assume the responsibilities of parenthood. The percentage of children born outside marriage more than doubled during the 1960s, from around 12 per cent to around 28 per cent.

Given the known dependency on time and prevailing conditions of genetic influences on fertility (Udry 1996; Kohler et al. 2006), we expected genetic influences to co-vary with environmental upheavals known to influence fertility decisions. We therefore based the interpretation of our findings on their variation across the five historical periods just identified.

Data and method

Data

The subjects of our sample were MZ and DZ twins who voluntarily participated in surveys of the TwinsUK registry. The sample comprises more than 12,000 individuals and has recorded 60,000 observations since 1992 (Moayyeri et al. 2013). The largest adult twin registry in the UK, it has been used primarily to study the genetic and environmental aetiology of age-related complex traits and diseases. The information it collects on participants includes some demographic data.

We aligned the design of our study with the earlier research, described above, on the early fertility of Danish twins and the number of their offspring. To allow more direct comparisons, we focused on cohorts born before 1968. Data on the AFB were collected by the main questionnaire and, from 2005, also by a behavioural questionnaire.

Respondents to the main questionnaire comprised a sample of 4,989 individuals. Respondents to the behavioural questionnaire added another 1,663. The combined sample of 6,652 was reduced by the removal of the following: 557 twin pairs because there were missing values in their demographic data; 223 twin pairs because they were males; 238 individuals because their zygosity was missing, 282 because their fertility data were missing, and 24 because they were not Caucasian. The removal of these cases resulted in a final sample of 4,548 individuals or 2,274 complete twin pairs. They were members of birth cohorts that ranged from 1919 to 1968 and had been interviewed in the period 1992–2010. The average age of the respondents used in the analysis was 57 years (Table 1), and most had passed beyond the reproductive period of their lives.

The variables used in our univariate twin models were AFB, age at censoring (C), zygosity, and year of birth. As is standard for survival models (Mills 2011), age was computed by subtracting the year when a respondent was born from the reported year at first childbirth or the last year of observation in the case of right censoring. For right-censored individuals, we used the age at last observation (ALO) as the censoring age C_i and replaced it with age 45 if it was higher. The age of 45 is the commonly assumed end of the reproductive lifespan and of the observation window for the fertility of females (Leridon 2008):

$$C_i = \begin{cases} ALO_i & \text{if } ALO_i \leq 45 \\ 45 & \text{if } ALO_i > 45 \end{cases} \quad (1)$$

Table 1 Summary statistics for female UK twins, birth cohorts 1919–68

Sample	Overall	MZ	DZ
N of twin pairs	2,274	41.6 (945)	58.4 (1,329)
Per cent censored individuals	16.8	18.2	15.8
Per cent twin pairs with censored information	27.1	28.1	26.4
Mean (SD) age	57.0 (10.5)	58.0 (10.6)	56.2 (10.3)
Mean (SD) age at censoring	55.4 (11.1)	55.0 (11.1)	55.7 (10.8)
Mean (SD) AFB	25.6 (4.7)	25.8 (4.6)	25.4 (4.7)
Correlation AFB (Pearson)	0.31	0.40	0.24

MZ, monozygotic; DZ, dizygotic; AFB, age at first birth; SD, standard deviation.
Note: Mean AFB and correlation AFB are computed for individuals who experienced AFB.
Source: TwinsUK.

Because only one twin had been interviewed in some waves of the survey, we used the life-history information to identify the most recent wave when information for both was available. Zygosity had been established by standard questions about physical similarity and confirmed by multiplex DNA genotyping in cases of uncertainty (Ooki and Asaka 2004).

Biosocial models and the twin design

The aim of behavioural genetics or biosocial models (Udry 1996) is to explain observed differences between individuals by differences in genetic and environmental factors. The use of models that require genetic information is rarely considered a requirement for demographic research (Mills and Tropf 2015), though the value of multidisciplinary and interdisciplinary studies in social science and demography is becoming more widely recognized (for an overview, see D'Onofrio and Lahey 2010; Balbo et al. 2013).

The most common way to disentangle the influence of genetic and environmental influences on a trait is to use twin data. Monozygotic (MZ) twins are genetically identical (i.e., share all their genotypes) but are assumed to have been exposed to common environmental influences, such as those of family and the neighbourhood in which they grew up. Fraternal or dizygotic (DZ) twins, in contrast, are akin to full siblings and share on average 50 per cent of their segregating genetic material. They are assumed to share their family environment to the same extent as MZ twins. Thus, the degree to which MZ twins are more similar in AFB than are DZ twins reflects the extent of genetic influence on AFB (for details, see Snieder et al. 1997, 2010; Boomsma et al. 2002).

A naive way of estimating the proportion of explained variance by additive genetic effects is to compute the correlations of a trait between siblings separately for MZ and DZ twin pairs. Since they are assumed to share family environment to the same extent and MZ twins are twice as similar as DZ twins in their genetic endowment, narrow sense heritability h^2 is twice the difference of the intra-group correlations of MZ and DZ (see Appendix A for a note on the concept of heritability). The effect of the shared environment of the twins is therefore the pairwise correlation of MZ minus h^2. Variance that is unexplained by these factors is due to non-shared environmental effects from outside or even within the family (Pike and Kretschmer 2009) and measurement error (for detail, see Snieder et al. 2010). The fitting of genetic models based on this logic has become standard in twin research.

At least three assumptions of the standard behavioural genetics 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), though evidence of it *not* being made is rare in studies (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. When the outcome of interest is fertility, this assumption might be considered rather strong, but if assortative mating is assumed to occur, the result is an underestimate of the heritability of genetic influences. The third assumption is that there are no non-additive genetic effects (see Appendix B). Previous studies found slight indications of dominance in their analyses (Kohler et al. 1999) as well as other gene–gene interactions (epistasis) for the physical ability to conceive a child (Christensen et al. 2003). To check for the presence of a significant dominant effect, we applied a standard behaviour genetics model-fitting approach to all of our models. The models estimating

dominant effects always provided a substantially lower fit with the data (details available on request).

Structural equation models

We fitted SEMs to estimate the influence of genetic and environmental factors on AFB. The underlying logic was in line with the comparison of intra-group correlations of twin pairs described above. However, these correlations have low power and large standard errors and do not make use of information available in variances and covariances (Snieder et al. 1997). SEM furthermore provides goodness-of-fit statistics thereby enabling us to test and compare alternative models (Snieder et al. 2010).

Figure 2 shows the ACE-path model. The capital letters in circles stand for the latent factors assumed to contribute to the observed variance in the sample. One-directional arrows refer to the directional non-standardized estimates of the respective variance components of the outcome: '*a*' represents additive genetic effects resulting from the sum of genetic effects of alleles (an allele is a variant of a gene for which different variants are possible) from all contributing loci (a loci is a location of a gene on a chromosome); '*c*' are environmental effects resulting from environmental influences shared between twins of a pair; and '*e*' are non-shared environmental effects resulting from the unique environment of an individual (including measurement error; for details, see Snieder et al. 1997, 2010). The model was constrained to be identical for each twin at the level of all parameters.

Bidirectional arrows indicate the assumed correlations, introduced earlier, between the latent variables for both groups of twins. The unique environment '*E*' is assumed to be independent for both types of twins. The boxes contain the measured outcome variable Y_j of the member $j \in \{1, 2\}$ of the twin pairs.

The basic structural equations in the model are

$$Y_1 = a * (A_1) + c * (C) + e * (E_1) \tag{2}$$

$$Y_2 = a * (A_2) + c * (C) + e * (E_2) \tag{3}$$

whereas

$$Cov\ (A_1, A_2) = \begin{cases} 0.5 & \textit{for DZ twin pairs} \\ 1 & \textit{for MZ twin pairs} \end{cases} \tag{4}$$

and

$$Cov\ (E_1, E_2) = 0 \tag{5}$$

To test how estimates differ between different kinds of censoring treatments, we estimated two models. First, we applied the classic SEM:

$$Y_j = AFB_j \tag{6}$$

where twin pairs with censored information do not contribute to the estimation. This has been the predominant or classic way of modelling these estimates in the literature (Neiss et al. 2002; Rodgers et al. 2008). However, as is known from alternative regression designs, '[b]inary and censored variables can lead to erroneous inferences about heritability in family studies […]' (Kohler and Rodgers 1999, p. 221). Another consequence is the reduction of the sample size. Also, this censoring treatment leaves it unclear whether the heritability patterns extend to twin pairs with censored information, which is an issue of particular interest given that the proportion of childless women in the UK doubled from 9 per cent for cohorts born in 1946 to 18 per cent for those born in 1958 (Kneale and Joshi 2008). Projections suggest that of the women born in 1970, from around 18 per cent (Bray 2007) to 22 per cent (Sobotka 2004) will still be childless at the age of 45. Childless women, therefore, are a group of growing demographic significance (Rendall and Smallwood 2003).

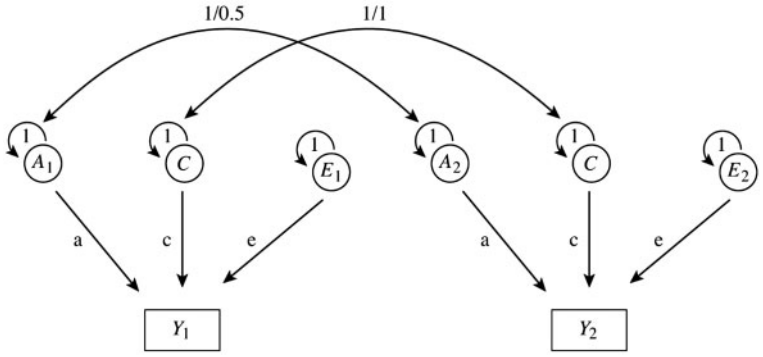


Figure 2 ACE-path model for the classic twin study (monozygotic/dizygotic twins)
Notes: A_1/A_2, genetic endowment; C, shared environment; E_1/E_2, unique environment; a, genetic effects; c, shared environmental effects; e, non-shared environmental effects/measurement error; Y_1/Y_2, outcome of Twin 1 and Twin 2.

In a second step we introduced a tobit model using the R-package *twinlm*, recently developed to estimate standard univariate twin models with default options and based on an R package for specifying and estimating linear latent variable models (Holst and Budtz-Jørgensen 2013). These are similar to other prominent open-access programs such as OpenMx (Boker et al. 2011), but *twinlm* in the *mets* package (Holst and Scheike 2015) allows the implementation of various statistical models for multivariate event history data such as the tobit model.

In the tobit model, the AFB_j^* of an individual i is assumed to be a latent, normally distributed trait which cannot be observed over its entire range. If a woman had been interviewed before having a child, we did not know at what point in her future she would have one, whereas the end of the reproductive lifespan at age 45 is the end of the observation window for all individuals. We therefore observed:

$$Y_j = \begin{cases} AFB_j^* & if\ AFB_j^* \leq C_j \\ C_j & if\ AFB_j^* > C_j \end{cases} \quad (7)$$

The maximum likelihood estimation for censored cases was based on the probability of being censored given the observed values (see also Kohler and Rodgers 1999). To standardize the estimates and report heritability and shared and non-shared environmental effects/measurement error, we divided a^2, c^2, and e^2 by the overall variance of AFB and denote genetic effects h^2. The variance components in the twin model were constrained to be between 0 and 1, as were the confidence intervals. To test for the significance of the genetic factor, we conducted a likelihood ratio test on the nested models of the ACE by constraining the genetic component to be zero, which is standard in behavioural genetics. The resulting p-values refer to the null hypothesis that there is no significant loss in model fit if we assume genetic effects to be zero.

Historical cohort analysis

For the comparison of cohorts historically, we followed two strategies. The aim of the first was to describe the underlying trend of changes in the genetic and environmental influences on AFB continuously. We therefore applied the tobit ACE model to twins from each birth year separately. For these estimates, we used locally weighted scatterplot smoothing (*Lowess*, Cleveland 1979) for the ten closest birth cohorts, which allowed us to estimate a focal value for each birth cohort. Local regression estimates were plotted in a smoothed scatterplot of a non-parametric regression of year born on the regression estimates to draw the most complete picture of historical changes in genetic and environmental influence on AFB (for details, see Appendix B).

Second, we allocated all individuals in the sample to the decades in which they were born. The oldest subsample contained the birth cohorts born in the period 1919–29, and the youngest those born in the period 1960–68. We applied both the classic and the tobit models to all subsamples in order to compare estimates. For the local linear regression we also added the confidence intervals for the subsample to the smoothed plot of the latent trend.

Results

Descriptive findings

Table 1 summarizes the variables for the overall sample (n = 2,274) as well as for MZ (n = 945) and DZ (n = 1,329) groups separately. The mean AFB of the overall sample is 25.6 years with a standard deviation of 4.7. It is slightly (0.4 years) lower for DZ than for MZ twins. The percentage of individuals who have not experienced AFB is 16.8 per cent, and information is censored for 27.1 per cent of twin pairs. The level of childlessness of MZ twins is 2.4 percentage points higher than that of DZ twins and the former are on average around 2 years older than the DZ twins. The mean AFB of the overall sample follows the population trend in AFB of the UK during the past century quite well at a slightly lower level (see the supplementary material for Appendix C).

Recall that one of our primary research questions asked to what extent individual differences in AFB could be explained by genetic factors. Table 1 reports a correlation in AFB of 0.24 for DZ twin pairs—for uncensored pairs only. DZ twins have the same degree of genetic relatedness as parents and children, and thus this value is in line with findings in Murphy's (1999) review of family correlations in fertility. The classic correlation in AFB of MZ twin pairs exceeds the correlation of DZ twin pairs in each sample of cohorts. This indicates a genetic component in AFB.

Twin models

Table 2 presents the results of the ACE model for the two SEM specifications: the tobit model, which adjusts for censoring, and the classic SEM model. Both models report significant genetic effects on the

Table 2 Standardized parameter estimates of ACE models in a classic SEM and a tobit model. UK cohorts born 1923–68

Cohort	Model	h^2	95 per cent CI	*p*-value	c^2	95 per cent CI	e^2	95 per cent CI	*n* (MZ-/DZ-pairs)
1919–68	(1 tobit)	0.26	[0.16–0.39]	<0.01	0.14	[0.07–0.25]	0.60	[0.55–0.64]	945/1,329
	(1 classic)	0.35	[0.21–0.52]	<0.01	0.07	[0.01–0.31]	0.58	[0.5–0.64]	773/1,119

h^2, heritability; c^2, shared environmental effects; e^2, unique environmental effect; 95 per cent CI, 95 per cent confidence interval; *n*, number of twin pairs; MZ, monozygotic; DZ, dizygotic.
Note: All models control for the year the twins are born.
Source: As for Table 1.

AFB of females in the UK born in the period 1919–68. Further, these effects exceed those of the shared environment and therefore are the main cause of within-family correlations in AFB. The tobit model yields an estimated heritability in AFB of 0.26. In other words, 26 per cent of the observed variance in AFB is explained by genetic variance. Variance in the shared environment of the twins (between-family variance) explains 14 per cent of the observed variance. The largest source of variance is attributable to the non-shared environment/measurement error, which accounts for 60 per cent of the observed variance in AFB.

Figure 3 shows the scatterplot of the smoothed non-parametric regression curves for the fitted values of the local linear regressions. This exploratory analysis of the data reveals a strong peak in genetic influence for cohorts born at the end of the 1930s. At the same time, the influence of the non-shared environment/measurement error experiences a strong drop. It furthermore depicts the continued growth of genetic influence for cohorts born from the end of the 1950s onwards. Whereas the influence of the shared environment is relatively stable, the non-shared environment/measurement error appears to decrease steadily, while it consistently remains the most important cause of variation.

In the most recent cohorts (1961–68), we observe a new pattern with a moderate level of heritability (0.32), and, compared with the previous cohort, an almost constant level of shared environmental influence (0.26). The non-shared environment/measurement error explains the highest share of the variation in all models, but is the lowest in the most recent cohorts, in which both the shared environment and the genetic influence have a moderate level of influence.

Figure 4 examines the heritability trend in more detail. It adds confidence intervals based on ACE models for the separate decades of birth cohorts shown in Table 3, and shows how the trend in each successive period is related to a significant event in the period. It can be seen that heritability estimates

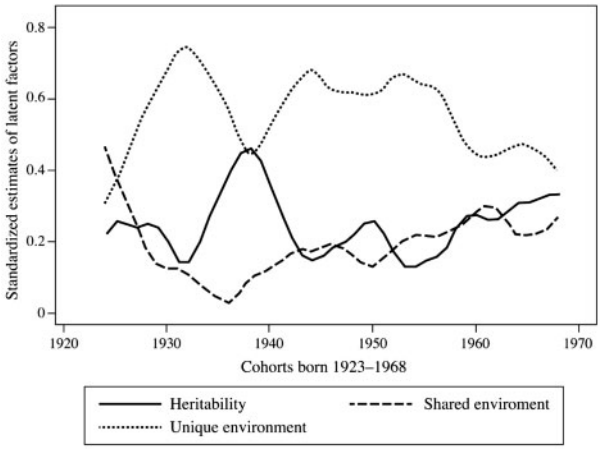


Figure 3 Smoothed scatterplot of local polynomial regressions of the birth year on the latent factor estimates for UK birth cohorts 1923–68
Notes: Smoothed results of local linear regressions based on the ten closest birth cohorts, UK cohorts born 1923–68.
Source: As for Table 1.

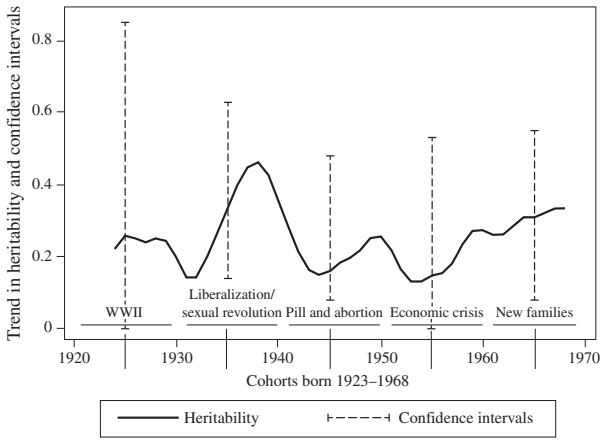


Figure 4 Smoothed scatterplot of local polynomial regressions of the birth year on heritability estimates for birth cohorts 1923–68 and the historical context
Notes: Smoothed results of local linear regressions based on the ten closest birth cohorts. Confidence intervals for decades are based on estimates in Table 3. UK cohorts born 1923–68.
Source: As for Table 1.

Table 3 Standardized parameter estimates of ACE models in a classic SEM and a tobit model for 10-year birth cohorts, UK cohorts born 1923–68

Cohort	Model	h^2	95 per cent CI	*p*-value	c^2	95 per cent CI	e^2	95 per cent CI	*n* (MZ-/DZ-pairs)
1919–29	(1 tobit)	0.20	[0.00–0.85]	0.56	0.30	[0.03–0.78]	0.50	[0.31–0.70]	40/43
	(1 classic)	0.51	[0.05–0.96]	0.06	0.00	[0.00–0.70]	0.49	[0.28–0.69]	30/30
1930–39	(2 tobit)	0.35	[0.14–0.63]	0.01	0.00	[0.00–1.00]	0.65	[0.55–0.74]	231/245
	(2 classic)	0.37	[0.12–0.69]	0.02	0.00	[0.00–1.00]	0.63	[0.51–0.73]	183/180
1940–49	(3 tobit)	0.23	[0.08–0.48]	0.09	0.16	[0.05–0.36]	0.61	[0.53–0.68]	293/444
	(3 classic)	0.43	[0.21–0.67]	0.00	0.00	[0.00–1.00]	0.57	[0.48–0.65]	220/345
1950–59	(4 tobit)	0.10	[0.00–0.53]	0.53	0.22	[0.08–0.43]	0.68	[0.58–0.77]	243/406
	(4 classic)	0.14	[0.01–0.59]	0.38	0.13	[0.01–0.46]	0.73	[0.61–0.83]	166/294
1960–68	(5 tobit)	0.26	[0.08–0.55]	0.10	0.32	[0.15–0.55]	0.42	[0.33–0.51]	138/191
	(5 classic)	0.05	[0.00–0.97]	0.79	0.38	[0.16–0.65]	0.57	[0.44–0.70]	80/130

h^2, heritability; c^2, shared environmental effects; e^2, unique environmental effect; 95 per cent CI, 95 per cent confidence interval; *n*, number of twin pairs; MZ, monozygotic; DZ, dizygotic.
Note: All models control for the year the twins are born.
Source: As for Table 1.

for cohorts born during the Second World War are not significant, as is confirmed by likelihood ratio tests provided in Table 3. The rise and peak in heritability are attained by cohorts born during the 1930s, which started childbearing in a period of liberalization and the sexual revolution. This trend reverses for cohorts of the 1940s, when contraception and legal abortion were introduced, and for cohorts born in the 1950s, who grew up during the economic crises. For these cohorts, heritability estimates once again become non-significant. Finally, there is a steady rise in heritability for women who became parents at the end of the 1950s, a period that saw a relaxation of normative pressure on familial behaviour and an increase in the popularity of cohabitation.

Please note that the standardized estimates provided are ratios of the three variance components divided by the total variance, and therefore the estimates shown can also vary with the overall variances. In view of this, we also applied the local linear regressions to the non-standardized genetic component. This showed that our results were robust to change in the overall variances (see the supplementary material for Appendix D).

Table 3 shows the ACE-model estimates for both the classic and the tobit models for all decades separately, illustrating the strong differences between the estimates of the two model specifications. In the overall sample (Table 2), the results for the standard model overestimate heritability by 9 percentage points. Comparing the results across cohorts, we see that there are even notable differences between cohorts of comparable sample size, such as the cohorts of the 1940s and 1950s. The estimates are also not consistent in their direction. While the classic model overestimates heritability of AFB in the first four cohorts, the estimate for the youngest one is lower than the estimate yielded by the tobit model and is non-significant.

Table 3 also provides the sample sizes of the submodels, showing first for the tobit model that they are particularly small for the oldest (MZ-pairs = 40, DZ-pairs = 43) and youngest cohorts (MZ-pairs = 138, DZ-pairs = 191), and second that they vary between model specifications. The implication of the small sample sizes is mainly that the non-significant heritability for the oldest cohorts might partly be the effect of power issues. Cohorts born in the 1950s, however, show non-significant results despite relatively larger sample sizes. In order to examine the model differences and the effect of the exclusion of censored twin pairs, we estimated all models in a classic SEM, imputing the C_i (see Appendix E). Results are in line with the tobit model, with slightly larger confidence intervals, suggesting that the imputation of the last age of observation remains a robust solution.

Summary and discussion

Our study investigated the relatively understudied topic of the influence of genes and the latent family environment on the AFB, and specifically on their influence on cohorts of UK twins born in the period 1919–68. The findings indicate that a substantial part —40 per cent—of the variation in AFB is caused by latent family influences, 26 per cent by genetic factors, and 14 per cent by the shared environment of siblings. The strongest influence—60 per cent—is that of the non-shared environment/measurement error, an influence to which the characteristics of the partner are likely to make a major contribution

(Kohler and Rodgers 2003). These findings suggest that genetic effects on reproductive behaviour are too strong to be ignored in demographic and sociological research on fertility.

There are two main ways in which genetic effects may influence human fertility. First, there can be a direct effect on such physiological characteristics as fecundity. For example, the age at which a female experiences menarche is strongly influenced by genetic variants (Anderson et al. 2007). This is also the age at which a woman is able to become pregnant. Thus, depending on the extent to which genes directly influence AFB, social scientists who do not take genetic effects into account actually underestimate the contributions of their predictors to the explanation of socially driven variance in reproductive behaviour (Kohler and Rodgers 2003).

Second, genetic influences may affect the processes of decision-making and life-course planning both consciously and subconsciously (Kohler et al. 2006). In fact, biologists now suspect that behavioural and psychological traits that are subject to genetic influences are those most closely linked to fertility, or more generally, to reproductive fitness (Kirk et al. 2001). Evidence of the importance of genetic influences on behaviour includes evidence of their effect on the age at first attempt to get pregnant (Kohler and Rodgers 1999; Rodgers et al. 2001). This suggests that part of the substantial genetic influence on the AFB is mediated by psychological traits.

Fertility behaviour is often seen to depend on the extent to which factors affecting the life course are experienced as certain or uncertain (Oppenheimer et al. 1997; Mills and Blossfeld 2005). The state of the social environment also appears to affect the extent to which biological predispositions affect the processes of decision-making (Udry 1996; Rodgers et al. 2007). Thus, research that has revealed large changes in the extent of genetic influences on fertility suggests that the main reason for the changes has been variation in the extent to which people have felt free to choose whether and when to become a parent and in what familial circumstances. In our study, economic uncertainty appears to override freedom from normative constraints to encourage the activation of genetic effects on decisions about childbearing. In particular, economic crises or exogenous shocks, such as those posed by war, are environmental conditions that may override individual differences in genetic influences on attitudes to childbearing. In contrast, freedom of choice in family formation, sexual behaviour, and birth control in modern societies are believed to encourage the activation of genetic influences on individual fertility behaviour (for details, see Udry 1996; Rodgers et al. 2007). That is why, as a second step in our study, we examined trends in the latent factor estimates for birth cohorts born in the period 1919–68. This revealed the pattern predicted by the foregoing considerations.

For women who entered adulthood during the Second World War, when major environmental constraints forced postponement of childbearing, we find no significant genetic influence on the AFB. In contrast, for individuals born in the 1930s, who benefited from liberalization of cultural values and changing sexual norms, heritability estimates show a strong peak. This pattern reverses for cohorts born in the 1940s, who started using effective contraception (the pill) and legal abortion. According to Hobcraft (1996), the fall in period fertility between 1964 and 1977 can probably be attributed largely to a reduction in unwanted births, but also to normative constraints on births outside marriage. For the first time, women were enabled by contraception to control their own fertility and avoid childbearing when it was normatively inappropriate, thus making it possible for environmental influences to override and suppress genetic influences on fertility. And in fact their influence becomes non-significant for cohorts born in the 1950s, which faced severe economic disruptions during the 1970s and the beginning of the 1980s—a period of economic crisis with a historical low point in total fertility (Hobcraft 1996).

For cohorts born in the 1960s we observe a new pattern: an increase in genetic influence but with a still relatively moderate level of family influence—mirroring a trend revealed in an earlier study of genetic influences on fertility behaviour (Kohler et al. 1999). These cohorts experience a new freedom in family and career planning (Van de Kaa 1987; Lesthaeghe 1995). Rodgers et al. (2008) maintain that the influence of stable environmental conditions after the Second Demographic Transition has had a continuing effect on women's career planning and education, in ways that have competed with their fertility planning.

Overall, our results emphasize the importance of the study of gene–environment interaction and its relationship with fertility behaviour. The results are consistent with the findings of Kohler et al. (1999) for Denmark, which showed a peak in heritability for completed fertility in the second half of the past century. The pattern we found thus appears to be robust for completed fertility and AFB in Denmark and the UK. A recent study by Bras et al. (2013)

also supports the reasoning underlying our interpretation of the findings of these studies. That study replicated the increase in heritability during the First Demographic Transition, using a database containing details of 100,000 sibling pairs born in the period 1810–70 in the Dutch Province of Zeeland. Similarly, a study by Rodgers et al. (2007), using data from the NLSY, shows that different genes are important throughout the life course, emphasizing the need for a life-course approach in behavioural genetics.

Although our theoretical reasoning provides a basis for the interpretation of our results, we do not provide a statistical test of gene–environment interaction. The differences in heritability estimates are not significant across birth cohorts and therefore remain suggestive. Replication is required, as are improvements in method. We believe that ongoing genome-wide data collection and complementary developments in methods of analysis (Yang et al. 2010, 2011; Tropf et al. 2015) will provide new insights and opportunities to address questions of gene–environment interaction directly and at the micro level.

Previous research studying the genetics of AFB has produced mixed results. While Kirk et al. (2001) reported findings broadly in line with our estimates from the tobit model, research from the US (Neiss et al. 2002; Rodgers et al. 2008) and Denmark found no significant genetic effect. We argued that this might be because these studies also used different analytical strategies. Kirk et al. (2001) applied an ordered probit model with individuals censored at age of last observation and age groups categorized to follow a normally distributed latent variable. Neiss et al. (2002) excluded childless observations (21 per cent) and Rodgers et al. (2008) included only twin pairs in which both siblings had experienced a first birth, and applied a classic linear SEM.

To demonstrate how different analytical strategies affect outcome, we applied two kinds of model: a recently developed tobit model in the SEM framework and the classic SEM that excludes censored cases for the covariance estimation. We found large differences between the two approaches, with different directions and degrees of deviation. We concluded that the choice of analytical strategy is critical. In particular, the exclusion of censored information appears to bias the estimate. We also applied classic models that imputed censored cases and found that they produced virtually the same results as the tobit model.

It is essential to note that fertility outcomes differ not only in consequence of the modelling strategy adopted, but are also dependent on the demographic and genetic characteristics included. Including only individuals who experience a first birth in the models gives a precise measure of AFB by focusing on a subset of the population that has had at least one child. But there are costs: important information in the data is not used, and the generalization of the results is restricted to a subpopulation. Epidemiological research that has sought genetic influences on other complex traits, such as blood pressure, supports the idea that subjects with extreme but unknown values of an outcome or phenotype (e.g., the use of antihypertensives) should not simply be excluded, because this leads to a bias in the components of familial variance and loss of power to identify genes (Cui et al. 2003; Tobin et al. 2005). Especially when comparing populations that differ in their extreme values, such as birth cohorts with different levels of childlessness, it is desirable to include all individuals in order to have comparable samples.

Our findings have important implications for future research. First, the two measures of AFB are genuinely different. As the outcomes for both demonstrate, there is no clear directional pattern in the results, and therefore no general conclusions can be drawn on how the inclusion or exclusion of childless individuals might affect results. For reasons given above, we believe that future studies should include a measure that includes childless individuals.

Second, censoring alone cannot entirely explain the mixed findings in the literature. First, in contrast to the studies by Neiss et al. (2002) and Rodgers et al. (2008), in our study the exclusion of childless individuals does not lead only to lower, but also to higher heritability estimates in subsamples. Second, the study by Nisén et al. (2013), which also excluded childless couples, reports a moderate level of heritability in Finland. We conclude that, while our study demonstrates how results can depend on the treatment of censored cases within one population, it is probable that gene–environment interaction plays an additional role in explaining the mixed findings across populations.

In conclusion, we draw attention to two substantive issues that merit attention in future research on the genetic study of fertility traits. First, it has been suggested that genes not only interact with social and psychological factors, but are mediated by them (Rodgers et al. 2001). We think future research should extend efforts to find effective genetic variants and to understand their function in relation to fertility behaviour. Second, the association between AFB and completed fertility needs to be

studied in more detail. Kirk et al. (2001) found a robust genetic covariation between AFB and completed fertility for Australian twins: genes predicting early AFB are associated with a higher number of offspring. Tropf et al. (2015) confirm these findings, using molecular genetic data on samples from the Netherlands and the UK. As Milot et al. (2011) conclude in a study of a population of natural fertility, AFB decreased within the past 300 years as a response to natural selection. In our view, multivariate genetic models need to be applied on large samples to determine the degree of genetic covariance in AFB and completed fertility and to better understand and predict the timing of human fertility.

Notes

1 Felix Christian Tropf is at the Department of Sociology, University of Groningen/ICS, email: ftropf@rug.nl. Nicola Barban and Melinda Mills are at the Department of Sociology, and Nuffield College, University of Oxford. Harold Snieder is at the Department of Epidemiology, University of Groningen and the University Medical Center, Groningen. Jornt J. Mandemakers is at the University of Groningen/ICS and the University of Wageningen.

2 The research leading to these results was funded by the Dutch Science Foundation (VIDI Innovation Grant 452-10-012 to M. Mills) and by the European Research Council Consolidator Grant SOCIOGENOME (to M. Mills, ERC-2013-CoG-615603). The authors gratefully acknowledge Tomas Sobotka for information and advice on age at first birth data for the UK. The TwinsUK study was funded by the Wellcome Trust and the European Community's Seventh Framework Programme (FP7/2007-2013). The study also received support from the National Institute for Health Research (NIHR)-funded BioResource, Clinical Research Facility and Biomedical Research Centre based at Guy's and St Thomas' NHS Foundation Trust in partnership with King's College London. They also thank the journal's anonymous reviewers for their helpful comments.

References

Anderson, C. A., D. L. Duffy, N. G. Martin, and P. M. Visscher. 2007. Estimation of variance components for age at menarche in twin families, *Behaviour Genetics* 37 (5): 668–677.

Balbo, N., F. C. Billari, and M. Mills. 2013. Fertility in advanced societies: a review of research, *European Journal of Population* 29(1): 1–38.

Boker, S., M. Neale, H. Maes, M. Wilde, M. Spiegel, T. Brick, S. Jefferey, R. Estabrook, S. Kenny, T. Bates, P. Mehta, and J. Fox. 2011. OpenMx: an open source extended structural equation modeling framework, *Psychometrika* 76(2): 306–317.

Boomsma, D., A. Busjahn, and L. Peltonen. 2002. Classical twin studies and beyond, *Nature Reviews Genetics* 3(11): 872–882.

Booth, A. L. and J. H. Kee. 2009. Intergenerational transmission of fertility patterns, *Oxford Bulletin of Economics and Statistics* 71(2): 183–208.

Bras, H., J. Van Bavel, and K. Mandemakers. 2013. Unraveling the intergenerational transmission of fertility: genetic and shared-environment effects during the demographic transition in the Netherlands, 1810–1910, *The History of the Family* 18(2): 116–134.

Bray, Helen. 2007. 2006-based national population projections for the UK and constituent countries, *Population Trends* 131: 8–18.

Christensen, K., H. P. Kohler, O. Basso, J. Olsen, J. W. Vaupel, and J. L. Rodgers. 2003. The correlation of fecundability among twins: evidence of a genetic effect on fertility? *Epidemiology* 14(1): 60–64.

Cleveland, W. S. 1979. Robust locally weighted regression and smoothing scatterplots, *Journal of the American Statistical Association* 74(368): 829–836.

Conley, D., E. Rauscher, C. Dawes, P. K. Magnusson, and M. L. Siegal. 2013. Heritability and the equal environments assumption: evidence from multiple samples of misclassified twins, *Behavior Genetics* 43(5): 415–426

Cui, J. S., J. L. Hopper, and S. B. Harrap. 2003. Antihypertensive treatments obscure familial contributions to blood pressure variation, *Hypertension* 41(2): 207–210.

Daw, J. and G. Guo. 2011. The influence of three genes on whether adolescents use contraception, USA 1994–2002, *Population Studies* 65(3): 253–271.

D'Onofrio, B. M. and B. B. Lahey. 2010. Biosocial influences on the family: a decade review, *Journal of Marriage and Family* 72(3): 762–782.

Dunne, M., N. G. Martin, D. Statham, W. Slutske, S. Dinwiddie, K. Bucholz, P. A. F. Madden, and A. C. Heath. 1997. Genetic and environmental contributions to variance in age at first sexual intercourse, *Psychological Science* 8(3): 211.

Felson, J. 2014. What can we learn from twin studies? A comprehensive evaluation of the equal environments assumption, *Social Science Research* 43: 184–199.

Fisher, R. A. 1930. *The Genetical Theory of Natural Selection*. London: Oxford University Press.

Freese, J. 2008. Genetics and the social science explanation of individual outcomes, *American Journal of Sociology* 114(S1): 1–35.

Goldstein, J. R., T. Sobotka, and A. Jasilioniene. 2009. The end of "Lowest-Low" fertility? *Population and Development Review* 35(4): 663–699.

Guo, G. 2005. Twin studies: what can they tell us about nature and nurture? *Contexts* 4(3): 43–47.

Guo, G., Y. Tong, and T. Cai. 2008. Gene by social-context interactions for number of sexual partners among white male youths: genetics-informed sociology, *American Journal of Sociology* 114(Suppl): 36–66.

Harden, K. P. 2013. Genetic influences on adolescent sexual behavior: why genes matter for environmentally oriented researchers, *Psychological Bulletin* 140(2): 434–465.

Hobcraft, J. 1996. Fertility in England and Wales: a fifty-year perspective, *Population Studies* 50(3): 485–524.

Hobcraft, J. 2006. The ABC of demographic behaviour: how the interplays of alleles, brains, and contexts over the life course should shape research aimed at understanding population processes, *Population Studies* 60 (2): 153–187.

Holst, K. K. and E. Budtz-Jørgensen. 2013. Linear latent variable models: the lava-package, *Computational Statistics* 28(4): 1385–1452.

Holst, K. K. and T. Scheike 2015. *Analysis of Multivariate Event Times*. R Foundation for Statistical Computing. Available: http://cran.r-project.org/web/packages/mets/mets.pdf.

Horwitz, A. V., T. M. Videon, M. F. Schmitz, and D. Davis. 2003. Rethinking twins and environments: possible social sources for assumed genetic influences in twin research, *Journal of Health and Social Behaviour* 44(2): 111–129.

Kiernan, K. 2004. Unmarried cohabitation and parenthood in Britain and Europe, *Law and Policy* 26(1): 33–55.

Kirk, K. M., S. P. Blomberg, D. L. Duffy, A. C. Heath, I. P. F. Owens, and N. G. Martin. 2001. Natural selection and quantitative genetics of life-history traits in western women: a twin study, *Evolution* 55(2): 423–435.

Kneale, D. and H. Joshi. 2008. Postponement and childlessness—Evidence from two British cohorts, *Demographic Research* 19(58): 1935–1968.

Kohler, H. and J. L. Rodgers. 2003. Education, fertility and heritability: explaining a paradox, in K. W. Wachter and R. A. Bulatao (eds.), *Offspring: Human Fertility Behaviour in Biodemographic Perspective*. Washington, DC: The National Academic Press, pp. 46–90.

Kohler, H. and J. L. Rodgers. 1999. DF-like analyses of binary, ordered, and censored variables using probit and tobit approaches, *Behaviour Genetics* 29(4): 221–232.

Kohler, H. P., F. C. Billari, and J. A. Ortega. 2002a. The emergence of lowest-low fertility in Europe during the 1990s, *Population and Development Review* 28(4): 641–680.

Kohler, H. P., J. L. Rodgers, and K. Christensen. 1999. Is fertility behaviour in our genes? Findings from a Danish twin study, *Population and Development Review* 25(2): 253–288.

Kohler, H. P., J. L. Rodgers, and K. Christensen. 2002b. Between nurture and nature: the shifting determinants of female fertility in Danish twin cohorts, *Biodemography and Social Biology* 49(3–4): 218–248.

Kohler, H. P., J. L. Rodgers, W. B. Miller, A. Skytthe, and K. Christensen. 2006. Bio-social determinants of fertility, *International Journal of Andrology* 29(1): 46–53.

Leridon, H. 2008. A new estimate of permanent sterility by age: sterility defined as the inability to conceive, *Population Studies* 62(1): 15–24.

Lesthaeghe, R. 1995. The second demographic transition in Western countries: An interpretation, in K. O. Mason and A.-M. Jensen (eds.), *Gender and family change in industrialized countries*. Oxford: Clarendon, pp. 17–62.

Lesthaeghe, R. and D. K. van de Kaa. 1986. Twee Demografische Transities? [Two demographic transitions?], in D. J. van de Kaa and R. Lesthaeghe (eds.), *Bevolking: Groei en Krimp*. Deventer: Van Loghum Slaterus, pp. 9–24.

Mealey, L. and N. L. Segal. 1993. Heritable and environmental variables affect reproduction-related behaviours, but not ultimate reproductive success, *Personality and Individual Differences* 14(6): 783–794.

Mills, M. 2011. *Introducing Survival and Event History Models*. London: Sage.

Mills, M. and F. Tropf. 2015. The biodemography of fertility: A review and future research frontiers, Kölner Zeitschrift für Soziologie und Sozialpsychologie. doi: 10.1007/s11577-015-0319-4

Mills, M. and H. P. Blossfeld. 2005. Globalization, uncertainty and the early life course: a theoretical framework, in H. P. Blossfeld, E. Klijzing, M. Mills, and K. Kurz (eds.), *Globalization, Uncertainty and Youth in Society*. London/New York: Routledge, pp. 1–24.

Mills, M., R. R. Rindfuss, P. McDonald, and E. te Velde. 2011. Why do people postpone parenthood? Reasons and social policy incentives, *Human Reproduction Update* 17(6): 848–860.

Milot, E., F. M. Mayer, D. H. Nussey, M. Boisvert, F. Pelletier, and D. Réale. 2011. Evidence for evolution in response to natural selection in a contemporary human population, *Proceedings of the National Academy of Sciences* 108(41): 17040–17045.

Moayyeri, A., C. J. Hammond, A. M. Valdes, and T. D. Spector. 2013. Cohort profile: Twins UK and healthy ageing twin study, *International Journal of Epidemiology* 42(1): 76–85.

Murphy, M. 1993. The contraceptive pill and women's employment as factors in fertility change in Britain 1963–1980: a challenge to the conventional view, *Population Studies* 47(2): 221–243.

Murphy, M. 1999. Is the relationship between fertility of parents and children really weak? *Biodemography and Social Biology* 46(1–2): 122–145.

Neiss, M., D. C. Rowe, and J. L. Rodgers. 2002. Does education mediate the relationship between IQ and age of first birth? A behavioural genetic analysis, *Journal of Biosocial Science* 34(2): 259–276.

Nisén, J., P. Martikainen, J. Kaprio, and K. Silventoinen. 2013. Educational differences in completed fertility: a behavioural genetic study of Finnish male and female twins, *Demography* 50(4): 1399–1420.

Ooki, S. and A. Asaka. 2004. Zygosity diagnosis in young twins by questionnaire for twins' mothers and twins' self-reports, *Twin Research* 7(1): 5–12.

Oppenheimer, V. K., J. Kalmijn, and N. Lim. 1997. Men's career development and marriage timing during a period of rising inequality, *Demography* 34(3): 311–330.

Pike, A. and T. Kretschmer. 2009. Shared versus non-shared effects: parenting and children's adjustment, *International Journal of Developmental Science* 3(2): 115–130.

Rendall, M. S. and S. Smallwood. 2003. Higher qualifications, first-birth timing and further childbearing in England and Wales, *Population Trends* 111: 18–26.

Rodgers, J. L., D. E. Bard, and W. B. Miller. 2007. Multivariate Cholesky models of human female fertility patterns in the NLSY, *Behavior Genetics* 37(2): 345–361.

Rodgers, J. L., H. P. Kohler, K. O. Kyvik, and K. Christensen. 2001. Behaviour genetic modeling of human fertility: findings from a contemporary Danish twin study, *Demography* 38(1): 29–42.

Rodgers, J. L., H. P. Kohler, M. McGue, J. R. Behrman, I. Petersen, P. Bingley and K. Christensen. 2008. Education and cognitive ability as direct, mediating, or spurious influences on female age at first birth: behaviour genetic models fit to Danish twin data, *American Journal of Sociology* 114(Suppl): 202–232.

Tobin, M. D., N. A. Sheehan, K. J. Scurrah, and P. R. Burton. 2005. Adjusting for treatment effects in studies of quantitative traits: Antihypertensive therapy and systolic blood pressure, *Statistics in Medicine* 24(19): 2911–2935.

Tropf, F., G. Stulp, N. Barban, P. Visscher, J. Yang, H. Snieder and M. C. Mills. 2015. Human fertility, molecular genetics, and natural selection in modern societies, *PLoS One*. Available: http://dx.plos.org/10.1371/journal.pone.0126821 (accessed: 11 June 2015).

Snieder, H., D. I. Boomsma, L. J. P. Doornen, and E. C. O. J. C. Geus. 1997. Heritability of respiratory sinus arrhythmia: dependency on task and respiration rate, *Psychophysiology* 34(3): 317–328.

Snieder, H., X. Xiaoling Wang, and A. J. MacGregor. 2010. *Twin Methodology*. Available: http://onlinelibrary.wiley.com/doi/10.1002/9780470015902.a0005421.pub2/abstract;jsessionid=874DC35AA37D05C3C7544F3381EB9308.f02t02?deniedAccessCustomisedMessage=&userIsAuthenticated=false.

Sobotka, T. 2004. *Postponement of Childbearing and Low Fertility in Europe*. Amsterdam: Dutch University Press.

Udry, J. R. 1996. Biosocial models of low-fertility societies, *Population and Development Review* 22: 325–336.

Van de Kaa, D. J. 1987. Europe's second demographic transition, *Population Bulletin* 42(1): 1–64.

Yang, J., B. Benyamin, B. P. McEvoy, S. Gordon, A. K. Henders, D. R. Nyholt, P. A. Madden, A. C. Heath, N. G. Martin, G. W. Montgomery, M. E. Goddard, and P. M. Visscher. 2010. Common SNPs explain a large proportion of the heritability for human height, *Nature Genetics* 42(7), 565–569.

Yang, J., S. H. Lee, M. E. Goddard, and P. M. Visscher. 2011. GCTA: a tool for genome-wide complex trait analysis, *The American Journal of Human Genetics* 88 (1), 76–82.

Appendix A: Heritability

The genetic component underlying a trait is named heritability and is the proportion of the genetically caused variance divided by the overall variance of the trait in a specific population. The genetic component can be further differentiated into additive and non-additive effects. For most quantitative traits—i.e., a trait that shows continuous variation—it is generally assumed that non-linear effects play only a minor role. In the fertility literature, evidence for gene–gene interaction is negligible. Therefore, the genetic component underlying a trait is commonly quantified as 'narrow sense' heritability and regarded as capturing all additive genetic effects (Mills and Tropf 2015).

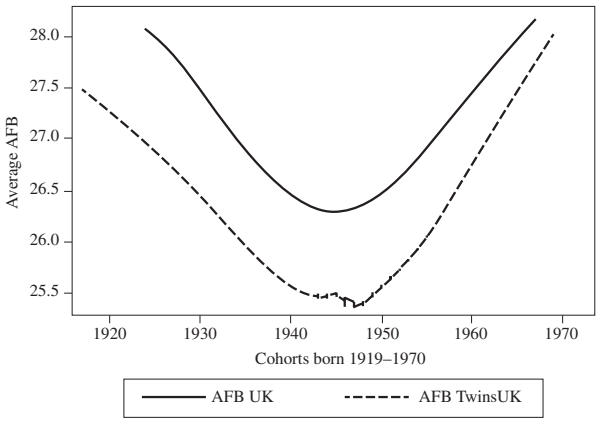
Appendix B: Local polynomial regression and non-parametric regression curve

In order to fully explore the underlying trends of the latent factors, we applied a local polynomial regression to the factor estimates of each birth cohort using the Stata package *Lowess*. We first estimated the ACE model of twin pairs of each birth year separately to produce time series of 46 estimates (birth cohorts born 1923–68) for the genetic influence, the shared environment, and the non-shared environment/measurement error. For cohorts born in the period 1919–22, the number of cases was too small to allow estimation.

In a second step, we partitioned the data into local areas around each birth year and, separately, fitted bivariate linear regressions of birth year onto the latent factors within each area. In practice, the choice of the local window of neighbourhood values included in the local regression is made by the researcher. It is a trade-off

between bias in the estimation and the remaining variance between the local estimates. As default value of the nearest neighbourhood method, *Lowess* uses 80 of the observations in the time series for each local estimation. We chose a lower value to reduce bias and increase the visibility of changes in latent factors. We provided the results for 22 per cent (equivalent to 10-year intervals) of the nearest neighbours, to show a good trade-off of reduced bias and acceptable variance in the factor trends. Note that the basic trends we discussed are not very sensitive and therefore the choice of this parameter was not critical for our substantive points.

Appendix C: Scatterplot of estimated AFB for birth cohorts 1923–68 from the UK[1] and TwinsUK[2]

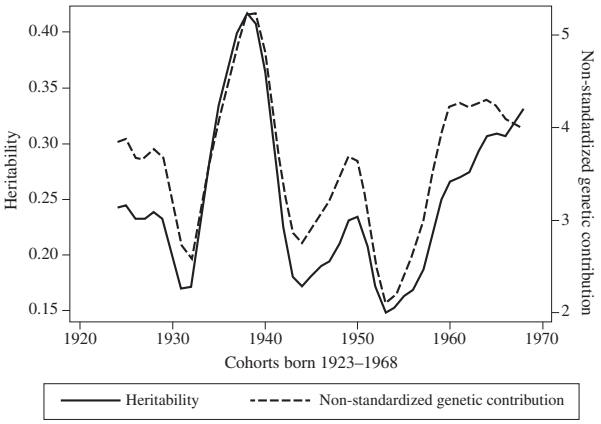


[1]Estimated average AFB in the UK. Because, historically, official data on birth order have been only collected within marriage, these values are based on estimates from the Office for National Statistics, Cohort fertility, Table 2. Available at: http://www.ons.gov.uk/ons/publications/re-reference-tables.html?edition=tcm%3A77-263133(accessed March 2013).

[2]Estimates are based on average AFB per birth cohort and smoothed in a linear local regression (see Appendix B) using 20 per cent of the closest neighbourhood values.

Sources: AFB UK and AFB TwinsUK.

Appendix D: Smoothed scatterplot of local polynomial regressions of the birth year on heritability (*y-axis*[1]) and the non-standardized genetic contribution (*y-axis*[2]). UK cohorts born 1923–68



Notes: Smoothed results of local linear regressions based on the ten closest birth cohorts. Confidence intervals for decades are based on estimates in Table 3. Cohorts born 1923–68: for birth cohorts.

Source: As for Table 1.

Appendix E: Standardized parameter estimates of ACE models in a classic SEM with imputed last year of observation in the case of censoring (Cim) and a tobit model for the overall sample and the cohorts, UK cohorts born 1923–68

Cohorts	Model	h^2	95 per cent CI	c^2	95 per cent CI	e^2	95 per cent CI	*n* (MZ-/DZ-pairs)
1919–29	(tobit)	0.20	[0.00–0.85]	0.30	[0.03–0.78]	0.50	[0.31–0.70]	40/43
	(Cim)	0.14	[0.00–0.96]	0.34	[0.04–0.84]	0.52	[0.30–0.73]	
1930–39	(tobit)	0.35	[0.14–0.63]	0.00	[0.00–1.00]	0.65	[0.55–0.74]	231/245
	(Cim)	0.36	[0.11–0.69]	0.00	[0.00–1.00]	0.64	[0.53–0.75]	
1940–49	(tobit)	0.23	[0.08–0.48]	0.16	[0.05–0.36]	0.61	[0.53–0.68]	293/444
	(Cim)	0.23	[0.06–0.53]	0.15	[0.03–0.41]	0.62	[0.52–0.70]	
1950–59	(tobit)	0.10	[0.00–0.53]	0.22	[0.08–0.43]	0.68	[0.58–0.77]	243/406
	(Cim)	0.10	[0.00–0.63]	0.20	[0.06–0.56]	0.70	[0.58–0.80]	
1960–68	(tobit)	0.26	[0.08–0.55]	0.32	[0.15–0.55]	0.42	[0.33–0.51]	138/191
	(Cim)	0.28	[0.08–0.60]	0.30	[0.11–0.56]	0.42	[0.33–0.53]	

Cim, classic model with imputed last year of observation in the case of censoring; h^2, heritability; c^2, shared environmental effects; e^2, unique environmental effect; 95 per cent CI, 95 per cent confidence interval; *n*, number of twin pairs; MZ, monozygotic; DZ, dizygotic. All models control for the year the twins are born.
Source: As for Table 1.