

University of Groningen

Social science genetics and fertility

Tropf, Felix

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version

Publisher's PDF, also known as Version of record

Publication date:

2016

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Tropf, F. (2016). *Social science genetics and fertility: Essays on the Interplay Between Genes, Social Environment and Human Fertility*. University of Groningen.

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

3. Human fertility, molecular genetics, and natural selection in modern societies⁷

Abstract: Research on genetic influences on human fertility outcomes such as number of children ever born (NEB) or the age at first childbirth (AFB) has been solely based on twin and family-designs that suffer from problematic assumptions and practical limitations. The present study exploits recent advances in the field of molecular genetics by applying genomic-relationship-matrix based restricted maximum likelihood (GREML) methods to quantify, for the first time, the extent to which common genetic variants influence the NEB and the AFB of women. Using data from the UK and the Netherlands (N = 6,758), we find significant additive genetic effects on both traits, explaining 10 percent (SE = 5) of the variance in the NEB and 14 percent (SE = 4) in the AFB. We further find a significant negative genetic correlation between AFB and NEB in the pooled sample of -0.62 (SE = 0.27, p-value = 0.02). This finding implies that individuals with genetic predispositions for an earlier AFB had a reproductive advantage and that natural selection operated not only in historical, but also in contemporary populations. The observed postponement in the AFB across the past century in Europe contrasts with these findings, suggesting evolutionary effects were overridden by environmental ones, and underscoring that evolutionary predictions in modern human societies cannot be straight forwardly tested. It emphasizes how an integrative research design from the fields of genetics and social sciences is needed to understand and predict fertility outcomes. Finally, our results suggest that we may be able to find genetic variants associated with human fertility when conducting GWAS-meta analyses with a sufficient sample size.

⁷

This chapter is based on Tropf, F. C., G. Stulp, N. Barban, P. M., J. Yang, H. Snieder, M. C. Mills (2015). Human Fertility, Molecular Genetics and Natural Selection in Modern Societies. *PLoS one*, 10(6), e0126821.

3. 1. Introduction

Recent research within both biology (Byars et al., 2010; Kirk et al., 2001; Milot et al., 2011) and demography (Kohler et al., 2006; Mills & Tropf, 2016) points to a genetic component of human fertility, namely the number of children ever born (NEB) and the age at first birth (AFB) of women, explaining up to 40-50 percent of the observed, respectively phenotypic variance in these traits. The well-established negative relationship of late AFB with lower NEB (Mills et al., 2011; Sobotka, 2004) appears to be partly genetic, suggesting that natural selection favored a younger age at first birth over the Twentieth century (Byars et al., 2010; Kirk et al., 2001; Milot et al., 2011). Genetic studies examining the relationship between NEB and AFB, however, have been solely based on twin (Kirk et al., 2001) or other family designs (Byars et al., 2010; Milot et al., 2011) that use data on expected genetic differences among relatives to estimate the genetic component underlying these traits. Although these studies are common in behavioural genetics, they can only draw indirect inferences about genetic contributions and suffer from problematic assumptions and practical limitations (critical discussions on, for example, the equal environment assumption (EEA) can be found in ref(Devlin, Daniels, & Roeder, 1997; Hettema, Neale, & Kendler, 1995; Horwitz et al., 2003). The studies are limited for further reasons. First, by virtue of their design, twin studies inherently require pairs of siblings and therefore exclude individuals from low fertility families, particularly only children, which may be problematic for the generalization of results. Second, dizygotic twinning has in contrast to monozygotic twinning a genetic base (Hoekstra et al., 2008; Mbarek et al., 2016; Montgomery, Zondervan, & Nyholt, 2014), which means that dizygotic twins potentially carry genes important for high fertility. Therefore, the use of monozygotic and dizygotic twins to investigate fertility questions in the classic twin design leads to a non-random genetic stratification and might bias variance estimates. Finally, a practical limitation of family designs is that they require data from multiple family-members, which are obviously more difficult to gather than data on unrelated individuals.

An ideal design to examine the genetics of fertility would be a direct estimate using single nucleotide polymorphisms (SNPs) across the entire genome for unrelated individuals who do not share the same micro environment, which was

first applied to height as a model complex trait (Visscher et al., 2010; Yang et al., 2010). The current study exploits recent advances in the field of molecular and quantitative genetics by applying genomic-relationship-matrix restricted maximum likelihood (GREML) methods to quantify for the first time the extent to which common genetic variants influence both the NEB and the AFB of women. We applied both uni- and bivariate models to these traits, producing unbiased estimates of their common SNP heritability and the extent to which the association between earlier AFB and higher lifetime fertility (NEB) is due to a (negative) genetic correlation between AFB and NEB (S. H. Lee et al., 2012). This not only helps us to understand the relationship between the AFB and NEB, but also allows an assessment of whether genes are associated with a reproductive advantage, indicating natural selection in contemporary, industrialized populations.

In contrast to twin and family designs, the GREML approach is free of confounding from shared environmental effects between close relatives because the method can be applied in a sample of unrelated individuals (Visscher et al., 2010; Yang et al., 2010). GREML analyses make use of the genetic similarity between pairwise unrelated individuals as captured by all common SNPs and correlate the genetic similarity with the phenotypic similarity between individuals (see material and methods). To ensure accurate and well-powered estimates, particularly for the bivariate model (Visscher et al., 2014), we pooled data sources to estimate the genetic influence on all outcomes of interest (see material and methods). We utilize two large cohorts, one from the Netherlands (NL, N = 4,338) and one from the United Kingdom (UK, N = 2,420, for descriptive statistics see Table 3-1). In both populations, resemblance in fertility outcomes has been reported for relatives (Murphy, 1999; Rijken & Liefbroer, 2009; Steenhof & Liefbroer, 2008) using intergenerational comparisons with survey data. However, no distinction between genetic and environmental effects responsible for this pattern could be made so far. After quality control of the merged genetic data files, we used more than 1 million SNPs to estimate the genetic relationships among the individuals (see material and methods) and subsequently the genetic variance components.

The most successful and popular design to detect the approximate location of genetic variants associated with a complex trait is the meta-analyses of genome-wide association studies (GWAS) from multiple samples. In lieu of this, our assessment of

the genetic effects of common SNPs based on the pooled samples shape the expectations to find individual variants when conducting a GWAS. We account for population stratification effects by adjusting for the first 20 principal components in our GREML models. Population stratification refers to allele frequency differences due to systematic ancestry differences. Population stratification can cause spurious associations if not adjusted properly (Price et al., 2006). We furthermore correct for country and birth cohort effects as well as dizygotic twinning. From the twin data only singletons are included, so that close relatives do not contribute to the estimates.

This study has several important implications for research in demography, genetics and biology. We know surprisingly little about genetic effects on human fertility at a population level, yet it is crucial for our understanding of fertility, the interpretation of related social science research in this field (Benjamin et al., 2012; Kohler & Rodgers, 2003; Rietveld, Cesarini, et al., 2013; Rietveld, Medland, et al., 2013; Rodgers et al., 2008) and broader questions of modern human evolution (Byars et al., 2010; Milot et al., 2011; Stearns et al., 2010; Zietsch et al., 2014). We first discuss the importance of adopting an integrative multidisciplinary approach to understand human fertility before proceeding with an introduction of the methods and the presentation and discussion of our findings.

3. 2 Towards an integrative approach in human fertility research

The term ‘fertility’ takes on different meanings in demography, reproductive medicine and biology (Mills et al., 2011). In demography, fertility refers to performance, specifically the two interrelated aspects of the tempo of childbearing (in our case age at first childbirth, AFB) and the quantum or number of children ever born (NEB) in a certain period (J Bongaarts & Feeney, 2000). In reproductive medicine, fertility defines the ability/inability of couples to conceive and have children given unprotected intercourse (Joffe, 2010). In biology, AFB and NEB have become central indicators for individual fitness as the successful transmission of genes to the next generation in post-industrial societies (Byars et al., 2010; Stearns et al., 2010), with particularly NEB shown to be nearly perfectly correlation with

alternative measures (Goodman, Koupil, & Lawson, 2012; Kirk et al., 2001). Due to improvements in hygiene and the reduction in prenatal, infant and child mortality in industrialized societies, NEB has emerged as the standard to measure lifetime reproductive success—that is, biological fitness (Stearns et al., 2010).

In the last decades, industrialized societies have experienced massive changes in both the postponement of AFB and drop in the total number of offspring, which cannot mainly be attributed to genetic or biological factors (Balbo et al., 2013; Mills et al., 2011). Rather, human reproduction is influenced by three analytically distinct but empirically interrelated factors: 1) genetic and biological fecundity (i.e., length of reproductive period, infertility diseases), 2) the environment (i.e., institutional and family structures); and, 3) reproductive choice of individuals (i.e., planned behavior, latent individual and partner characteristics).

Previous research has successfully demonstrated that there is a genetic component to reproduction with over 70 genome-wide association studies (GWAS) published for 32 traits and diseases associated with reproduction (Montgomery et al., 2014). This includes identification of genes such as those related to age at menarche (Liu et al., 2009; Sulem et al., 2009), menopause (He et al., 2009; Perry et al., 2009; Stolk et al., 2009, 2012) and endometriosis (Painter et al., 2011). Environmental factors, such as women's gains in education and labor market participation, gender equity and economic uncertainty have been shown to strongly impact the tempo and quantum of fertility (for reviews see Balbo et al., 2013; Mills et al., 2011). Studies of reproductive choice, which have examined the predictive power of fertility intentions on behaviour, often position reproductive choice in a socio-psychological framework that consists of attitudes (perceived costs and benefits), norms (influence social network) and perception of control over individual choice (Ajzen, 1991; W B Miller & Pasta, 1994).

A bivariate twin model in a study by Rodgers and colleagues (Rodgers, Hughes, et al., 2001) suggests an interrelation between reproductive choice and genetic factors, providing evidence for shared genetic effects on the decision to have a first child and the number of children during one's lifetime. It is therefore likely not only that biological fecundity, the environment and reproductive choice interact with each other, but also that genes influence reproductive choice. Genetic endowment in social science fertility research has been virtually ignored (Balbo et

al., 2013), yet may be of major importance when interpreting the observed associations.

If the quantum of fertility in the form of NEB is at least partly genetically influenced, this implies that certain SNPs have a higher chance of being successfully transmitted to the next generation than others, and by extension that the allele frequency might change due to natural selection—that is, that evolution may be occurring. If the negative relationship between AFB and NEB is partly genetic, this would indicate that the AFB was under natural selection during the Twentieth century and that more recent birth cohorts may carry a higher genetic predisposition for an earlier AFB.

Using a family-design, findings from the Framingham Heart Study demonstrated that the same genes influencing NEB are negatively correlated with the AFB (Stearns et al., 2010). The authors subsequently predict that selective changes in the disposition for the timing of the first child predict the decrease in the AFB for subsequent generations. The study design, however, is based on correlations between relatives and the estimates can therefore be inflated by shared environmental factors such as family norms that are important for fertility (van de Kaa, 1987). Family designs cannot robustly discriminate between the case where the correlation between NEB and AFB is environmentally caused, and natural selection, where the correlation is genetically caused and the allele frequencies of the genome might change (Stearns et al., 2010). This limitation leaves a less desirable practical solution: "...to note the issue and remain modest in drawing conclusions" (Stearns et al., 2010, p. 613). In the current study, our design does make it possible to directly draw conclusions about contemporary natural selection based on information derived from the field of molecular genetics. When the trait of interest, here the age at first birth, does not genetically covary with fertility, a genetic response to selection will not occur (Morrissey, Kruuk, & Wilson, 2010).

3. 3 Data & method

3. 3. 1 Data

For the Netherlands, we use data from LifeLines, a multi-disciplinary prospective population-based cohort study examining, in a unique three-generation

design, the health and health-related behaviours of 167,729 persons living in the North of The Netherlands, including genotype information from more than 13,000 unrelated individuals (Klijs et al., 2015). It employs a broad range of investigative procedures in assessing the biomedical, socio-demographic, behavioural, physical and psychological factors which contribute to the health and disease of the general population, with a special focus on multi-morbidity and complex genetics.

For the UK, we use data from TwinsUK, the largest adult twin registry in the country with more than 12,000 respondents (Moayyeri et al., 2013). Due to our analytical strategy, we randomly selected only one twin for analysis and controlled for dizygotic twinning as a genetically related process. We recognize that for generalizability a population-based sample such as LifeLines is more desirable for the models we present. The descriptive statistics of the phenotypic variables in the genotyped subsamples with full fertility information are shown in Table 3-1.

3.3.2 Genotypes

Since genotyping had been performed using different chips in the UK and the Netherlands, we use imputed data to aid the alignment of both datasets. The HapMap3 imputation panel has been shown to be reliable for GREML analysis (S. H. Lee, Wray, Goddard, & Visscher, 2011).

We received genotype data from TwinsUK and Lifelines, which we imputed according to the 1000 genome panel, after which we selected HapMap3 SNPs with an imputation score larger than 0.6. For quality control (QC), we excluded the SNPs with a larger missing rate than 1 percent, lower minor allele frequency than 1 percent and which failed the Hardy-Weinberg equilibrium test at a threshold of 10^{-6} for both datasets. We merged the TwinsUK and the Lifelines samples and quality controlled the merged dataset in the same way again. On average 1,017,420 SNPs could be utilized to estimate the GRM between individuals. We used the software Plink (Purcell et al., 2007) for the quality control and merging of the two datasets.

3. 3. 3 Phenotypes

Number of children ever born

Number of children ever born measures the number of children a woman has given birth to including stillbirths. This has been asked directly in the twinsUK (“How many children have you given birth to?”) or we constructed it using questions about the year of childbirth of each child. In Lifelines, respondents have been asked to list the birth and death date of children from their current and previous partner with up to 6 children in both categories. For the Lifelines and part of the TwinsUK questionnaires, information for the date of death of the children was given. In both datasets less than 0.2 percent of the children had not reached reproductive age and the correlation of number of children ever born and number of children reaching reproductive age was >0.98 .

Since fertility is strongly age dependent, we focus on women with completed fertility history in reference to the phenotype. In general, the end of the woman’s reproductive lifespan occurs around the age of 45 (Leridon, 2008). Thus, we only included women aged 45 or older in our analysis of NEB. Furthermore, in vitro fertilization (IVF) – often related to twinning and multiple births – can bias results if IVF compensates for genetically based infertility. However, in our TwinsUK sample, only 60 women reported using IVF, and these were excluded in the final analyses.

Age at first birth

To calculate the AFB, we used information on the year of childbirth of the first child and year of birth of the mother. In TwinsUK, information from an additional behavioral questionnaire directly asking for the age at first birth in 2005 was available. Childless individuals were treated as missing in the analysis.

3. 3. 4. Heritability estimates

The genetic component underlying a trait is commonly quantified in terms of heritability (h^2) as the proportion of the genetically caused variance (σ_G^2) over the overall phenotypic variance of the trait (phenotype, V_p):

$$h^2 = \frac{(\sigma_G^2)}{(\sigma_p^2)}$$

Whereas the phenotypic variance is the sum of genetic and environmental σ_e^2 variance components.

$$\sigma_p^2 = \sigma_G^2 + \sigma_e^2$$

The methods we applied have been detailed elsewhere (Visscher et al., 2014, 2010; Yang, Lee, et al., 2011; Yang et al., 2010). Briefly, we applied a mixed linear model

$$y = g + e$$

where y is an $N \times 1$ vector of dependent variables, N is the sample size, g is the $N \times 1$ vector with each of its elements being the total genetic effect of all SNPs for an individual, and e is an $N \times 1$ vector of residuals.

We have $g \sim N(0, \sigma_G^2 A)$ and $e \sim N(0, \sigma_e^2 I)$, where σ_G^2 is the genetic variance by all SNPs, A is the genetic relationship matrix (GRM) estimated from SNPs, σ_e^2 is the residual variance and I is an identity matrix. The variance components are estimated using the restricted maximum likelihood (REML) approach. The NEB is not normally distributed (see Supporting Information 3-1). This might bias the inference, yet simulation studies show that there is no bias even for binary traits (S. H. Lee et al., 2011). Still, we base our p-values on likelihood-ratio tests, comparing the full model with one constraining genetic effects to be zero (Snijders, 2011).

The estimates of heritability obtained using GREML can be interpreted as the proportion of variance of a trait based on a large set of common genetic variants genotyped. The method is based on the genetic relatedness among individuals measured on about one million of SNPs.

This analysis has been extended to a bivariate approach by Lee and colleagues (2012) to estimate unbiased genetic correlation based on a standard bivariate linear mixed model combined with the genome-wide genetic relatedness matrix.

Genetic correlation

The genetic correlation ($r(G)$) is an estimate that standardizes the genetic covariance between two traits $Cov(G_{i_1,i_2})$ by the genetic variance of both traits:

$$r(G_{i_1,i_2}) = \frac{Cov(G_{i_1,i_2})}{\sqrt{V_{G_{i_1}}} * \sqrt{V_{G_{i_2}}}}$$

If the genetic correlation between two traits is 1, all genetic variance in trait 1 and 2 has a common origin. If the genetic correlation is 0, the genetically based variances of trait 1 and trait 2 are independent.

Phenotypic and genetic correlation analysis

The phenotypic correlation between two traits $r(P_{i_1,i_2})$ is the sum of genetic and environmental influences shared across traits and can be estimated like this:

$$r(P) = \sqrt{h_{i_1}^2} * r(G_{i_1,i_2}) * \sqrt{h_{i_2}^2} + \sqrt{e_{i_1}^2} * r(E_{i_1,i_2}) * \sqrt{e_{i_2}^2}$$

whereas h_i^2 is the heritability of trait i in the model and e_i^2 is the environmental or residual variance contribution for the trait, standardized for the overall variance

$$e^2 = \frac{\sigma_e^2}{\sigma_p^2} = 1 - h^2$$

And $r(E_{i_1,i_2})$ is the environmental or residual correlation between the traits (for the estimates of environmental effects see Supporting Information 3-4). We can solve this to compute the fraction of the phenotypic correlation explained by the genes (or the environment respectively the residuals). For the transformation of standard errors, the delta-method has been applied (Lynch & Walsh, 1998).

3. 4 Results

3.4.1 Descriptive findings

Table 3-1 shows the descriptive statistics for both traits in the TwinsUK and the Lifelines cohorts. Overall the AFB is around one year later in the Dutch (26.83) than in the British cohort (25.70), and the British women are about 9 years older. These characteristics are interrelated, since Europe experienced a massive postponement in the AFB during the second half of the Twenties century (Mills et al., 2011), so the larger proportion of younger individuals leads to a later average AFB in the data.

To combine the cohorts, both fertility measures, AFB and NEB, were standardized by country (Z-transformation) and the NEB was log transformed to approach normal distribution (see Supporting Information 3-1 for distributions and 3-2 for the model estimation of all alternative transformations - estimates are robust across transformations).

3.4.2 The correlation between AFB and NEB

In line with previous studies, women who had their first child at a later age had a lower number of children ever born (Fig. 1) in both the British and Dutch samples. The observed correlation for individuals with full information on both traits (therefore excluding all childless individuals, individuals younger than 45 and individuals without information about the AFB) between AFB and NEB is -0.32 (N=1,521) in the UK cohorts, -0.26 (N = 2,553) in the Dutch cohorts and -0.28 (N = 4,074) for the standardized measures in the pooled cohorts (-0.27 if estimated from the residuals of all covariates, not listed).

3.4.3 SNP heritability of AFB and NEB

Table 3-2 depicts the SNP based heritability (h^2_{SNP}) estimated from the univariate models for AFB and NEB. Both traits have a significant genetic component, with h^2_{SNP} for NEB of 0.10 (SE 0.05) and for the AFB of 0.14 (SE 0.04). These results suggest that additive effects of common SNPs explain 10 percent of the variance in the NEB and 14 percent of the variance in the AFB of women.

3.4.4 Bivariate GREML analysis of AFB and NEB

Table 3-3 shows the results of the bivariate GREML model for AFB and NEB, including the genetic correlation between both traits. The genetic correlation would be $-1.00/1.00$ if all genetic effects of AFB and NEB are shared, and 0 if the genetic effects of AFB and NEB are completely independent. The genetic correlation estimate is -0.62 (SE 0.27) and significantly different from 0 (p -value = 0.02), meaning that genes which lead to a later age at AFB are indeed negatively associated with the NEB. Based on these estimates, genetic effects lead to a phenotypic correlation of -0.07 (0.03) between AFB and NEB, whereas the overall correlation estimated from the fitted model is -0.38 (SE = 0.02). Therefore around 20 percent of the phenotypic correlation is attributable to shared genetic effects across the traits, while still the main part is associated with common environmental/residual effects of AFB and NEB. The phenotypic correlation estimated from the genetic model is larger than the observed correlation because the bivariate GREML analysis does not require both traits to be measured on exactly the same set of individuals; that is, it makes use of additional information, e.g. childless individual for the estimates of NEB. If we only include individuals with full information on both traits in the genetic model – as we do when computing the phenotypic correlation directly – the phenotypic correlation estimated based on the genetic model (-0.29 SE = 0.02) is not significantly different from the observed value of the Pearson correlation (-0.27), and the component due to genetic effects estimated from the GREML model (-0.08 SE = 0.05) is not significantly different from the estimate using all available information (-0.07 SE = 0.03), whereas the inference would be weaker (see Supporting Information 3-4 for the model excluding all individuals with missing information).

Table 3-1 Descriptive statistics of the female TwinsUK and Lifelines samples

	TwinsUK				Lifelines			
	Mean	SD	Min-Max	N	Mean	SD	Min-Max	N
Birth year	1951	13	1919-1987	2420	1960	11	1920-1989	4338
AFB	25.70	4.74	15-44	1951	26.83	4.26	16-43	4016
NEB	2.07	1.21	0-9	1990	2.25	1.20	0-9	2875

Note: N for the age at first birth (AFB) is different from the N for number of children ever born (NEB). The reason for this is that only women older than 45 have been included in the analysis of NEB. For example, a 35 years old woman with a first child is part of the analysis for AFB but not for NEB. Therefore in the Lifelines cohorts the N for AFB is larger than for NEB, because it contains a large proportion of women younger than 45. This also implies that the average AFB in the more recent birth cohorts is younger than in the overall population.

Source: Lifelines Cohort Study and TwinsUK

Table 3-2 Heritability estimates of NEB and AFB for the pooled sample of women from the UK and the Netherlands using information from about 1 million SNPs

	h^2_{SNP} (SE)	p -value ^c	N
Number of children ever born ^a	0.10 (0.05)	0.02	4865
Age at first birth ^b	0.15 (0.04)	0.0004	5967 ^d

Notes: a: standardized by country and log transformed to adapt the distribution; b: standardized by country; c: p-values are based on likelihood-ratio tests, the reference model constraints genetic effects to be 0; Please find estimates of untransformed variables in SI 3-3; d : The N for age at first birth is larger than for number of children ever born. The reason is that only women with completed fertility history are included for the latter (for discussion see material and methods and Table 3-1).

Source: Same as Table 3-1

3. 5 Discussion

Using recently developed analytical techniques from molecular genetics we provide direct evidence for a genetic component underlying the AFB and NEB of women in the UK and the NL born during the Twentieth century. Moreover, genetic effects on the tempo (AFB) and quantum (NEB) of human reproduction co-vary, which partly explains why women who start reproducing at an earlier age have higher fertility.

This genetic association between AFB and NEB can have different origins. Both traits might simultaneously be influenced by the same genetic effects (pleiotropy). Alternatively, genetic effects on the NEB could be mediated via the AFB, or some combination of the two. To further examine the causal relations between these factors, actual measured genotypes important for these traits could be integrated into the statistical models (van den Oord & Snieder, 2002) for applications such as Mendelian randomization (Verduijn, Siegerink, Jager, Zoccali, & Dekker, 2010) – although this has been challenged (Conley, 2009). Whatever the underlying cause of the genetic association between NEB and AFB, as the consequence of this genetic association we expect natural selection acting in modern, industrialized societies, implying that women born in more recent cohorts may be genetically

inclined to have an earlier AFB. This prediction of a decrease in AFB, however, is somewhat of a 'population paradox' because it strongly contradicts observed fertility trends over the last 50 years. Instead, there has been a massive postponement in the AFB of an average of 4-5 years in nearly all European countries and the US since the 1970s (Mills et al., 2011).

Although our results seem to generate a paradox, they are well in line with studies on natural fertility populations, such as from Milot and colleagues (2011), who observed a decrease in AFB as a response to natural selection in a contemporary population. One probable explanation is that natural selection works in addition to environmental forces and in the opposite direction, with the latter being far stronger. Natural fertility populations are assumed to have set fertility norms that maximize reproductive success. In the absence of contraception, full reproductive potential can be realised (Milot et al., 2011). In European and many industrialized societies, in contrast, environmental changes across the past century such as the use of contraceptives and women's educational expansion and entry into the labour market have had a strong impact on fertility behaviour (Mills et al., 2011; Sobotka, 2004). This which has led to a postponement in the AFB even though more recent populations in the Netherlands and the UK are genetically predisposed to an earlier AFB. In that sense, the environment has overridden the forces of natural selection.

The discrepancy between observed changes in fertility and those predicted by evolutionary processes show parallels to the case of human height. Although there is evidence for negative natural selection height at least in US populations (Byars et al., 2010; Stulp, Verhulst, Pollet, & Buunk, 2012), people still, on average, grow up to be taller than their parents (Komlos & Lauderdale, 2007). This is also largely attributed to environmental factors, such as better nutrition and improved health care (Steckel, 1995). Natural selection, however, may also work in tandem with environmental factors: a recent study suggests that, in the Netherlands, natural selection favoured taller heights, and thus reinforced the effects of improved environmental quality over the last 150 years (Stulp et al., 2015).

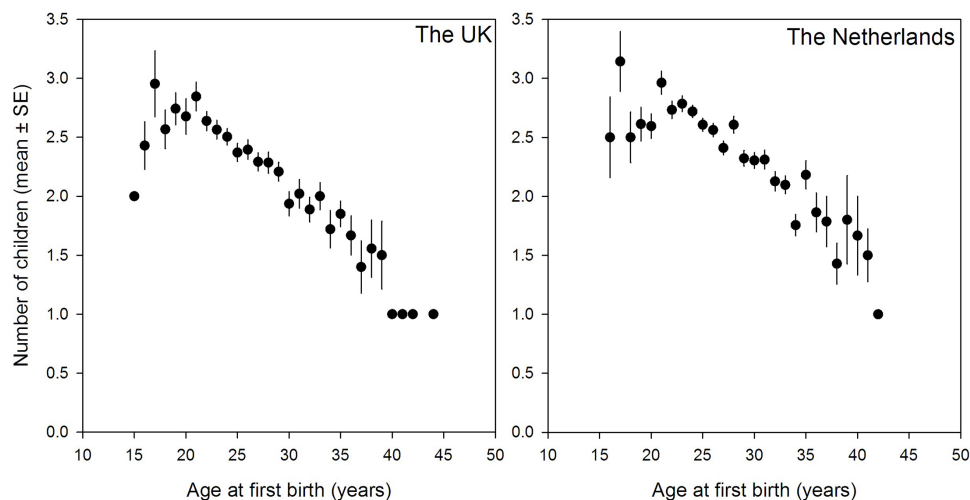


Figure 3-1 The association between age at first birth and number of children ever born in the British and the Dutch cohorts.

Source: Same as Table 3-1

A second potential – and largely related – explanation for the fact that AFB is postponed despite selection towards genes favouring earlier birth is that genes and the environment interact across birth cohorts. Previous twin studies have in fact shown differences in heritability estimates across cohorts and environments in both NEB (Kohler et al., 1999) and AFB (Kirk et al., 2001; Neiss et al., 2002; Nisén et al., 2013; Rodgers et al., 2008; Tropf, Barban, et al., 2015). Therefore, independent of additive environmental effects leading to postponement in AFB, genetic variants important for the AFB may differ across cohorts and populations, so that large changes due to natural selection do not necessarily occur.

The genetic effects estimated in this study represent narrow sense heritability estimated from SNP data. As would be expected (Vinkhuyzen, Wray, Yang, Goddard, & Visscher, 2013), they are lower than the estimates of narrow sense heritability (~0.20-0.30) obtained from family designs. Potential reasons for this are, on one hand, the inflation of estimates by shared environmental factors in family designs, and on the other hand, true genetic effects of variants that are not captured through linkage disequilibrium with SNPs used in GREML analysis. In order to

Table 3-3 Estimates of the bivariate genetic model for NEB and AFB for the pooled sample from the UK and the Netherlands using information from about 1 million SNPs

$h^2_{\text{SNPs}_{\text{NEB}}}(\text{SE})$	$h^2_{\text{SNPs}_{\text{AFB}}}(\text{SE})$	$r(\text{G})_{\text{SNPs}_{\text{AFB-NEB}}}(\text{SE})$	p-value ^a	Phenotypic correlation		$N_{\text{AFB/NEB}}$
				Overall (SE) ^b	Due to genetic effects (SE) ^b	
0.08 (0.05)	0.15 (0.04)	-0.62 (0.27)	0.02	-0.38 (0.02)	-0.07 (0.03)	5967 / 4865 ^c

Note: NEB: standardized by country and log transformed to adapt the distribution; AFB: standardized by country; a: p-values are based on likelihood-ratio tests, the reference model constrains genetic effects to be 0; – one-tailed (default in GCTA); b. Standard errors have been transformed using the delta method; c: The N of age at first birth is larger than for number of children ever born. The reason is that only women with completed fertility history are included for the latter (for discussion see material and methods and Table 3-1). For the full model, including environmental/residual effects see S3.

Source: same as in Table 3-1

more rigorously examine of genetic effects as well as gene-environment interplay, replication in larger datasets and across populations is required. The provision of data with genetic and environmental information continues to grow, as do more advanced analysis techniques (Zaitlen et al., 2013).

Nonetheless, it becomes obvious that human fertility is both a genuinely biological process as well as a social undertaking. We conclude from our findings that an integrative approach between the social and biological sciences is necessary to better understand the changing patterns in, and even to predict future, levels of human fertility.

Despite the significant advances in the estimation techniques and sample size of this study, there are two limitations that need to be made explicit. First, the interpretation of NEB in an evolutionary manner implies an interpretation of NEB as a *measure of fitness*. It would be better to have information on the number of children who entered reproductive age or even more appropriate, the number of grandchildren entering reproductive age; this would permit a more precise measure of how far genes have been successfully transmitted across generations. However, NEB has been shown to be a good measure of reproductive success (Stearns et al., 2010) due to diminishing mortality during the reproductive lifespan. Recent genetically informed research furthermore demonstrates that the same which influence NEB also influence the number of grandchildren born and therefore have a long-term effects (Zietsch et al., 2014). Second, in contrast to most research in demography (Mills, 2011), it is still uncommon to address cases right censored information (i.e., those who have not yet had a child by the time of observation) in genetic studies. In our study, we treated individuals who remained childless as missing when estimating genetic influences for the AFB, since they did not yet have a child. Childless individuals, however, are of great interest for demographic research as well as from an evolutionary perspective – since they are the ones who do not transmit their genes to the next generation. While the structural equation modelling in twin studies provides alternative solutions such as Tobit (see chapter 2 and Tropf, Barban, et al., 2015) or ordered models (Kirk et al., 2001) to integrate censored information, there remains no possibility of doing this in current applications of GREML. In general, the observed association can be expected to be stronger when including childless individuals, because all childless women are

right-censored cases. However, it has been shown in twin models that it is more difficult to predict the effect on the genetic analysis. Therefore, the extent of empirical differences between our model and survival models must be tackled in future research.

To date, thousands of genetic variants have now been successfully linked to physical or psychological traits (Hindorff et al., 2009; Visscher et al., 2012), as well as complex 'socio-genetic' traits like educational attainment (Rietveld, Medland, et al., 2013) and traits related to reproduction (Montgomery et al., 2014). We conclude that our study, based on the same genetic data as in GWAS studies, raises confidence that we will find genetic variants associated with human fertility when conducting GWAS-meta analyses of a sufficient sample size.

