Mega-analysis of 31,396 individuals from 6 countries uncovers strong gene-environment interaction for human fertility

Felix C. Tropf[a,b], Renske M. Verweij[a], Peter J. van der Most[c], Gert Stulp[d], Andrew Bakshi[e], Daniel A. Briley[f], Matthew Robinson[e], Anastasia Nyman[g], Tõnu Esko[h,i], Andres Metspalu[h], Sarah E. Medland[j], Nicholas G. Martin[j], Harold Snieder[c], S. Hong Lee [e,k], Melinda C. Mills[b]

[a]Department of Sociology/Interuniversity Center for Social Science Theory and Methodology, University of Groningen, Groningen 9712 TG, The Netherlands

[b]Department of Sociology/ Nuffield College, University of Oxford, Oxford OX1 3UQ, UK

[c]Department of Epidemiology, University of Groningen, University Medical Center Groningen, Groningen 9700 RB, Netherlands.

[d]Department of Population Health, London School of Hygiene and Tropical Medicine, London WC1E 7HT, UK

[e]The University of Queensland, Queensland Brain Institute, Brisbane, QLD 4072, Australia

[f] Department of Psychology, University of Illinois at Urbana-Champaign, Champaign 61820-9998, USA

[g]Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, PO Box 281, Stockholm SE-171 77, Sweden

[h]Estonian Genome Center, University of Tartu, Tartu, Estonia, 51010

[i]Broad Institute of the Massachusetts Institute of Technology and Harvard University, 140 Cambridge 02142, MA, USA

[j]Quantitative Genetics Laboratory, QIMR Berghofer Medical Research Institute, Brisbane, QLD 4029, Australia

[k]School of Environmental and Rural Science, The University of New England, Armidale NSW 2351, Australia

Corresponding author: Felix C. Tropf, Manor Rd, Ox1 3UQ, +44 (0) 1865 28 17 40, fctropf@gmail.com

Categories: Social Sciences, Biological Sciences

Keywords: Fertility; Age at first birth; Gene-environment interaction; Missing heritability; Natural selection; GREML

Abstract

Family and twin studies suggest that up to 50% of individual differences in human fertility within a population might be heritable. However, it remains unclear whether the genes associated with fertility outcomes such as number of children ever born (NEB) or age at first birth (AFB) are the same across geographical and historical environments. By not taking this into account, previous genetic studies implicitly assumed that the genetic effects are constant across time and space. We conduct a mega-analysis applying whole genome methods on 31,396 unrelated men and women from six Western countries. Across all individuals and environments, common single-nucleotide polymorphisms (SNPs) explained only ~4% of the variance in NEB and AFB. We then extend these models to test whether genetic effects are shared across different environments or unique to them. For individuals belonging to the same population and demographic cohort (born before or after the 20th century fertility decline), SNP-based heritability was almost five times higher at 22% for NEB and 19% for AFB. We also found no evidence suggesting that genetic effects on fertility are shared across time and space. Our findings imply that the environment strongly modifies genetic effects on the tempo and quantum of fertility, that potentially ongoing natural selection is heterogeneous across environments, and that gene-environment interactions may partly account for missing heritability in fertility. Future research needs to combine efforts from genetic research and from the social sciences to better understand human fertility.

Significance statements

Fertility behavior – such as age at first birth and number of children – varies strongly across historical time and geographical space. Yet, family and twin studies, which suggest that up to 50% of individual differences in fertility are heritable, implicitly assume that the genes important for fertility are the same across both time and space. Using molecular genetic data (SNPs) from over 30,000 unrelated individuals from six different countries, we show that different genes influence fertility in different time periods and different countries, and that the genetic effects consistently related to fertility are presumably small. The fact that genetic effects on fertility appear not to be universal could have tremendous implications for research in the area of reproductive medicine, social science and evolutionary biology alike.

Introduction

Twin and family studies from Western countries show that genetic factors may explain up to 50% of the differences in human fertility outcomes such as number of children ever born (NEB) or age at first birth (AFB) within a population (1–8). It remains unknown, however, whether the same genes are important for fertility across different environments or whether gene-environment interaction modifies genetic effects on fertility. This is a vital question for at least three reasons. First, the most successful and widely-used design to detect the approximate location of genetic variants associated with complex traits is a meta-analysis of genome-wide association studies (GWAS) from multiple populations (9). This approach assumes genetic effects on a trait to be universal across environments. However, concerning fertility, this requires investigation given that environmental upheavals such as the introduction of the pill or educational expansion have substantially changed fertility behavior in the recent past (10, 11). A second and interrelated point is that studies resorting to molecular genetic data to quantify heritability as the variance in a trait explained by genetic variance result in lower estimates than family studies (12) – and this is true also in fertility research (1, 2, 7, 13, 14). This discrepancy might, amongst other reasons, be a consequence of the interaction between environment and genes. Family studies are conducted amongst members of the same populations, whereas for example GWAS use data from individuals across populations. If genes can explain variance in fertility within but not between populations, heritability estimates based on different populations will be smaller than within populations (12, 15). Third, Fisher's fundamental theorem of natural selection predicts at environmental equilibrium (close to) zero additive genetic effects on fitness-related traits such as fertility, because genes that reduce fitness are expected to have been passed on to the next generation to a lesser extent (16). Nevertheless,

additive genetic influences on fertility are well established and a potential explanation is that the genes that are important for fertility differ across environments (17).

Twin and family designs cannot be used to answer the question as to whether different genes are important for fertility across populations or birth cohorts since relatives usually live in the same country and twins always have the same age. However, with the advent of molecular genetic data and complementary analytical techniques and software, it has become possible to examine the genetic material of unrelated individuals across different (historical) populations and therefore the unique possibility exists to test whether the same genes influence a trait across diverse environments (18–22). In this study, we exploit these advances for the first time, by empirically assessing whether genetic effects on fertility differ across geographical and historical environments.

We pooled a series of large datasets consisting of 31,396 unrelated (~ second cousin, IBS<0.05, see Material and Methods) genotyped men (n = 10,489) and women (n = 20,907) from six countries and seven study populations for analysis (for the US: HRS, ARIC; for the Netherlands: LifeLines; For Sweden: STR/SALT; for Australia: QIMR; for Estonia: EGCUT; for the UK: TwinsUK) who are assumed to have completed their reproductive period ($age_{men} > 50; age_{women} > 45$). We first conducted a mega-analysis, which is based on individual information from different populations in contrast to a meta-analysis that uses summary statistics of analyses conducted within populations, and applied whole genome methods (20, 21) using GCTA software (18) to estimate SNP-heritability (h^2_{SNP}). SNP-heritability is the proportion of total phenotypic variance that is explained by common genome-wide SNPs. Based on a previous study using data from women from the Netherlands and the UK, we expect h^2_{SNP} to be around 0.10 for number of children ever born and around 0.15 for AFB (23).

Second, to investigate gene-environment interaction, we follow two strategies: the first one consists in fitting multiple genetic relatedness matrices in our model, one global matrix for all individuals and more matrices indicating whether individuals lived in the same population and/or were part of the same birth cohort. The global matrix estimates the effects genes have across all environments, whereas the population/birth cohort specific matrices estimate context specific genetic effects (see 18 and Material and Methods for our specifications). The second strategy consists in fitting bivariate genetic models to investigate the moderating effect of the postponement transition on genetic effects on fertility (see 22, 24 and Material and Methods for our specifications). This model allows us to estimate h^2_{SNP} separately for different birth cohorts as well as the genetic correlation across them. To maximize power in these models, we divided all populations into two demographic birth cohorts. A central turning point in the reproductive environment of the 20th century occurred when AFB experienced a massive postponement of up to 4-5 years in nearly all advanced societies, the so-called 'postponement transition' (25), or Second Demographic Transition (10, 11, 26, 27). The primary reasons proposed for fertility postponement have been women's increased educational attainment and their employment in the labour force, triggered by factors such as the availability of effective contraception (10, 11). Cultural transformations in terms of sexual freedom, family planning and the timing and role of children are also central (26, 27). To investigate the moderating effect of fertility postponement we divide individuals into birth cohorts born either before or after this massive postponement in AFB in the past century (10, 11, 25, 28).

Results

Descriptive findings

The descriptive statistics for NEB and AFB for all populations under study (LifeLines, TwinsUK, STR, Estonia, HRS, ARIC and QIMR) as well as the pooled data separate for men

and women can be found in Supplementary Table S1. The participants were born between 1903 and 1967. The mean number of children per woman is 2.0 in Estonia, Sweden and the UK and 3.3 in Australia. For men, the lowest reported number of children is in Sweden and Estonia with around 1.9 children per man and is the highest in Australia at around 3.4. AFB was available for the Netherlands, UK, Sweden, Estonia and Australia. For both men and women, it was lowest in Estonia with an average of 24.6 for women and 27.7 for men and highest in Australia with 26.7 and 29.8 for men. Individuals who start reproducing at a later age have fewer children, with correlations between NEB and AFB ranging between -0.24 (Netherlands) and -0.38 (Australia; Supplementary Table S2). This pattern is less consistent across countries; for example in Australia, the highest fertility levels are observed, despite having the highest AFB. This reflects heterogeneity in fertility levels across countries with Australia having traditionally higher fertility levels than other Western countries (for a trend comparison of the total fertility rate across countries see Supplementary Figure S1).

Demographic Trends

Figure 1 shows the trends in AFB during the 20th century for the countries in our study based on population data if available (see Material and Methods for details). We observe the well-established U-shaped pattern of AFB of a falling AFB in the first half of the 20th century followed by a turning point and upturn in the trend of AFB towards older ages. This postponement transition in fertility timing was accompanied by a strong drop in completed fertility in most countries (29).

Sociocultural and technological changes, such as the introduction of effective contraception, educational expansion or changing norms in reference to sexuality and family planning, have largely driven these trends (10, 11). These environmental changes occurred in specific time periods in each country. In order to test for gene-environment interaction in our

analyses, we split the data into birth cohorts born before and after the turning point of fertility postponement to reduce environmental heterogeneity amongst the individuals who are members of the same birth cohort. This turning point differs across countries (Figure 1) with Australia having the earliest start of postponing (1939) and Estonia the latest (1962; see Supplementary Information Table S3 for all turning points and details). Differences in the onset of the postponement transition are well established and can be due to, for example, political reasons. This is the case of Estonia, for example, where early AFB had been strongly promoted by political incentives when it still was part of the Soviet Union in time periods prior to 1990 (30).

[Figure 1 here]

Genetic effects on fertility from the whole genome

Model 1: SNP heritability of AFB and NEB across environments

Not taking environmental differences into account, SNP based heritability (h^2_{SNP}) is significant and low for number of children ever born and age at first birth (Table 1). For NEB, h^2_{SNP} is 0.038 (SE = 0.0097, p-value = 2.0×10^{-5}) and for AFB it is 0.053 (SE = 0.019 p-value = 0.0020; these estimates are based on the full genetic relatedness matrix - see Material and Methods). These findings mean that around four per cent of the variance in NEB and around five per cent in AFB can be attributed to common, additive genetic effects in the pooled data. These estimates are much lower than those reported in other studies (23).

Model 2: Genes x population interaction (g x p)

A potential reason to explain why estimates are lower than expected is that the SNPs important for fertility have different effects across environments. Model 2 therefore adds an interaction term to Model 1 that captures the influence of genetic variance on fertility only within populations. The gene-population interaction models for NEB and AFB show that

shared genetic effects across populations are much lower than genetic effects within populations. With respect to NEB, the shared genetic effects across populations are negligible

[Table 1 here]

(0.0070, SE = 0.011, p-value = 0.26), whereas within populations additional additive effects are estimated to be 0.15 (SE = 0.024, p-value = 6.0×10^{-12}). The same applies to AFB, where shared genetic effects are estimated to be only 0.024 (SE = 0.022, p-value = 0.14), whereas the within population effect is 0.10 (SE = 0.039, p-value=0.0037; Table 1; Model 2). These results show that there is little overlap in SNPs that influence fertility across populations, and that most of the SNPs influencing fertility are population specific.

Model 3: Genes x demographic birth cohort (g x d)

Similar to the Model 2, in which we modeled population specific effects, we also examined whether there were genetic influences on fertility that were specific to birth cohorts. We find that there is additional genetic variance explanation for individuals who live in the same demographic cohort. While h^2_{SNP} for all birth cohorts is estimated at zero for both NEB (SE = 0.013, p-value = 0.50) and AFB (SE = 0.03, p-value = 0.35), for individuals living in the same demographic cohort there is a significant additional genetic variance component of 0.097 (SE = 0.017, p-value = 3.3×10^{-16}) for NEB and 0.084 (SE = 0.031, p-value = 4.6×10^{-4}) for AFB (Table 1; Model 3). Thus, similar to what we observed for the different populations, we find that SNPs influencing fertility traits are specific to cohort.

Model 4: Genes x population x demographic birth cohort (g x p x d)

Including a gene-environment interaction term that takes into account both the population and the demographic cohort simultaneously (Model 4), we observe that for NEB the interaction with demographic cohort (0.064, SE = 0.020, p-value = 5.9×10^{-4}) and the interaction with population and demographic cohorts (0.085, SE = 0.045, p-value = 0.0030) are significant.

This suggests that living in the same demographic cohort increases h^2_{SNP} independent of whether individuals live in the same population, but rather living in the same population and the same demographic period additionally increases h^2_{SNP}. For AFB h^2_{SNP} is only significantly different from zero for individuals living in the same population and demographic cohort (0.18, SE = 0.077, p-value = 0.0032).

Overall SNP based heritability for each model

Subsequently, overall heritability estimates were calculated as the sum of the different components of each model to examine the increase in heritability estimates when including the different interaction terms (See Figure 2 corresponding to Supplementary Table S4). The overall h^2_{SNP} for NEB increases almost fivefold, from 0.04 (SE = 0.01; Model 1) to 0.22 (SE = 0.026), when population and demographic cohort are taken into account. For AFB, the trend is very similar, with h^2_{SNP} of 0.053 (SE = 0.019) in the baseline Model 1 and 0.19 (SE = 0.039) in the genes x population x demographic cohort interaction model, when population and demographic cohort are taken into account.

[Figure 2 here]

Sensitivity analysis: Genes x Sex

The analyses presented are based on pooled datasets of men and women. However, two data sources contain (almost) only women (TwinsUK and ARIC). To the extent that different genes influence fertility across sexes, this might drive the observed differences across populations. We therefore conducted a sensitivity analysis extending Model 3 to a genes x population x sex interaction model. We find that considering sex-differences does not significantly improve the model fit (p-value for AFB 0.5, for NEB 0.093) and therefore are confident that our findings do not result from sex-differences (Supplementary Table S6).

Bivariate analysis

We complementarily estimated a bivariate model based on Model 2 and splitting data for demographic cohort (see Material and Methods), which allows us to estimate genetic effects across (σ_g^2) and within (σ_{gxp}^2) populations separately for different demographic birth cohorts and investigate whether genetic effects are correlated across demographic birth cohorts. Table 2 shows that σ_{gxp}^2 estimates for NEB within populations are significant for both demographic cohorts before ($\sigma_{gxp}^2/\sigma_p^2$ = 0.15, SE = 0.039, p-value = $9.6\text{x}10^{-6}$) and after ($\sigma_{gxp}^2/\sigma_p^2$ = 0.13, SE = 0.048, p-value = 0.0010) fertility postponement. It furthermore shows a positive correlation of genetic effects on NEB across demographic cohorts within populations (1.00, SE = 0.35, p-value = $1.3\text{x}10^{-5}$). In Model 4 of Table 1, this remained suggestive, since the genetic effects within populations but shared across demographic cohorts ($\sigma_{gxp}^2/\sigma_P^2$) were non-significant (0.073, SE = 0.036, p-value = 0.18).

[Table 2 here]

The bivariate model for the AFB finds some evidence that in both demographic cohorts genetic effects are observed (before fertility postponement 0.099, SE = 0.073, p-value = 0.083; after fertility postponement 0.13, SE = 0.074, p-value = 0.070), although these effects were marginally significant. However, there is no evidence that genetic effects correlate across demographic birth cohorts (0.11, SE = 0.59, p-value = 0.27), which is well in line with the null-estimate $\sigma_{gxp}^2/\sigma_P^2$ (0.00, SE = 0.062, p-value = 0.50).

Genetic effects shared across all populations (σ_g^2) are only significant for NEB and birth cohorts born before fertility postponement (σ_g^2/σ_P^2 = 0.031, SE = 0.018, p-value 0.042) while non-significant for younger cohorts and for AFB. Genetic correlations are therefore not interpreted.

Discussion

Using data from seven populations and six countries, we demonstrate that genetic effects on fertility outcomes –number of children ever born (NEB) and age at first birth (AFB) – differ across temporal and spatial environments. For NEB, genetic effects within populations are stronger than across populations, but correlate between individuals who were born before or after the turning point in fertility postponement of the 20th century. For AFB, genetic effects are only significant if individuals live in the same demographic cohort and the same population. The full gene-environment interaction model (Model 4) as well as the bivariate analyses provide no evidence for shared genetic effects for each phenotype across populations and demographic birth cohorts. Our results show that different SNPs are associated with fertility traits in different populations and birth cohorts, and there are hardly any genetic effects that are consistently related to these traits across populations and cohorts. Our results uncover a strong interplay of genetic and environmental factors influencing human fertility.

Quantitative geneticists have been puzzled by low heritability estimates based on GWAS findings or even whole-genome estimates such as GREML model as we apply it in the current study, describing the phenomenon of 'missing heritability' (12). Previous attempts to explain missing heritability partly by non-additive genetic effects remain empirically untested (31) or find only little support (32). Our findings of strong gene-environment interaction imply first that the detection of genetic variants associated with fertility traits is a major challenge using meta-analyses of GWAS on individuals from different populations. Likewise, predictions out of the discovery sample might be difficult, because discovered SNPs might have different effects in different samples. Second, they imply that lower heritability estimates from GWAS studies compared to GREML approaches or family studies might be due to the fact that genetic effects are (to some extent) not universal but context specific. In the model considering gene-

environment interaction across population and demographic cohort, we report heritability findings of 0.22 for NEB and 0.19 for AFB (see Figure 2 and Supplementary Information Table S4), which are fourfold higher than across all contexts and approach heritability estimates from family models (2, 14). It is therefore central to understand the cultural and environmental factors that interact with human fertility as well as their origins across (family) environments in order, for example, to define missing heritability or validate the findings from twin studies. It is to be noted that our findings are probably fostered by the strong behavioural and social nature of fertility, which might be more sensitive to cultural and societal heterogeneity than for example morphological traits. A recent investigation by Yang et al. (33) shows that missing heritability for the anthropometric traits height and body mass index is negligible when using whole genomic sequencing data in a new GREML model and assuming that family models overestimate heritability.

Demonstrating that genetic effects on fertility outcomes differ across environments, our study substantially contributes to the current knowledge on the genetic architecture of human reproduction. Previous twin studies show for several countries and birth cohorts that fertility outcomes such as NEB and AFB are genetically influenced (1, 2). However, it remained unclear whether the same genes are associated with fertility across environments. Using molecular genetic data and GREML methods (18, 20, 21), we were able to relate the genetic material of individuals across environments and found that common SNPs explain substantially more variance within than between countries and birth cohorts for fertility traits.

Previous twin and family studies furthermore suggest that the level of heritability of fertility traits can change across time and space (5, 7, 34, 35). However, these differences could not be statistically validated. In the current study, we proposed a multi-matrix approach to test for gene-environment interaction but also applied bivariate GREML models across birth cohorts (22, 24). Bivariate GREML models allow estimating SNP-heritability within two

independent samples was well as the genetic correlation across them. We cannot confirm the suggestion that the level of heritability changed over time, but find that heritability levels are comparable before and after the strong fertility postponement in the past century.

Different levels of heritability have also been reported across countries (2). Our multi-matrix GREML approach distinguishes between pairs of individuals who are living in the same or in different populations. The resulting within population estimate is therefore an average across all populations and we cannot compare different levels of heritability across populations. A more desirable study design would be a multivariate genetic modelling approach as we presented it in a bivariate design to investigate differences across demographic birth cohorts. However, this approach was not possible in our study due to small sample sizes in each population and a consequent lack of statistical power (22), but might become feasible in the future with better data availability.

Our findings are of interest to scientists within the medical, biological and social sciences alike (1, 2, 36, 37). Research has successfully identified genetic variants associated with reproductive diseases and traits (37). However, it remains unknown how these affect realized fertility. We find no evidence that genetic effects underlying fertility in one country predict fertility outcomes in another one. Genetic effects on fertility outcomes are rather strongly dependent on an individual's environment. Recently, social scientists have made large efforts to integrate molecular genetics into their research (1, 2, 23, 34, 38–42). However, when it comes to reproductive health, environmental factors are also likely to be critical in understanding how genetic factors are modified in relation to fecundity and infertility.

For evolutionary biologists, our findings have at least two important implications. First, the number of children ever born has been used as a proxy for fitness, given the diminishing child mortality rate in contemporary societies (4, 23, 36). Additive genetic variance therefore indicates natural selection under environmental equilibrium within populations: if all else

equal, genes that lead to a higher number of offspring will have a higher frequency in future generations. Due to natural selection, Fisher predicted additive genetic variance in fertility to be (close to) zero in the absence of gene environment interaction, since genes that reduce fitness are passed on to the next generation to a lesser extent thereby reducing their frequencies (16). Nevertheless, we find significant additive genetic influences on fitness traits such as NEB and AFB – substantial yet lower than heritabilities observed for morphological traits such as height (14, 15, 23, 43). Finding significant genetic influences on this these proxies of fitness suggests that, along with sociocultural changes surrounding fertility, genetic variants under selection have also changed (for review see 1, for review see 2, 5, 7, 17, 34, for comment see 44–46). Gene-environment interaction can explain why we find additive genetic variance in fitness related traits despite natural selection.

Second, previous research has uncovered an ongoing natural selection in contemporary societies (3, 4, 23, 47, 48) and even attempted to forecast changes in for example height and blood pressure across generations (4). For valid evolutionary predictions about observable changes in traits across generations due to natural selection, fertility needs to be consistently heritable, the same genes need to be under selection across generations and the direction of the selection needs be similar. Our results demonstrate moderate genetic influences on fertility within populations indicating potentially ongoing human evolution. However, this potential is delimited in at least two ways: First, genetic effects on fertility strongly differ across countries and therefore may lead to heterogeneity across human populations rather than to universal changes in humans. Second, the finding that genetic effects underlying proxies of fitness vary so markedly across time periods suggests that substantial caution is needed when inferring long-term evolutionary predictions.

For social scientists, genetic influences had been originally thought of as biological constraints on human reproductive behavior (42). Yet some previous studies showed that

genetic predispositions may underlie decision making processes on fertility timing and motivation (6, 7, 49, 50). It has been suspected that genetically based behavioural and psychological traits have become more important than physiological ones in the recent past (6, 8, 34, 51). This hypothesis remains to be tested, but our results confirm that genetic influences on fertility have evolved with social changes in the reproductive environment and therefore underscore the necessity to integrate social factors into genetic research on fertility.

Overall, our study uncovers great challenges for investigations into the genetic architecture of fertility, which can only be overcome by interdisciplinary work between both social scientists and geneticists using ever larger datasets, with combined information from genetics and social sciences (36).

Material & Methods

Cohorts

In this study we combined data from seven cohorts and six countries. For the US, we use data from ARIC, HRS, for Estonia from EGCUT, for Australia QIMR data from the Australia Twin and Family Register, for the Netherlands the LifeLines Cohort Study, for the United Kingdom TwinsUK and for Sweden the STR. All studies have received ethical approval.

ARIC

ARIC (Atherosclerosis Risk in Communities Study) is a community-based prospective cohort study of 15,792 adults, ages 45–64. Participants were identified by probability sampling from four U.S. communities (Forsyth County, North Carolina; Jackson, Mississippi; suburban Minneapolis, Minnesota; and Washington County, Maryland) and were enrolled between 1987 and 1989 (52–54).

HRS

The Health and Retirement Study (HRS) is an ongoing cohort study of Americans, with interview data collected biennially on demographics, health behavior, health status, employment, income and wealth, and insurance status. The first cohort was interviewed in 1992 and subsequently every two years, with 5 additional cohorts added between 1994 and 2010. The full details of the study are described in (55).

EGCUT

Estonian data come from of the Estonian Genome Center Biobank, University of Tartu (EGCUT, www.biobank.ee), a population-based database which comprises the health, genealogical and genome data of currently more than 51,530 individuals (56). Each participant filled out a Computer Assisted Personal Interview including personal data (place of birth, place(s) of living, nationality etc.), genealogical data (family history, three generations), educational and occupational history and lifestyle data (physical activity, dietary habits, smoking, alcohol consumption, and quality of life).

QIMR

Data for Australia was received from the Queensland Institute for Medical Research (QIMR). The participants were drawn from cohorts of adult twin families that have taken part in a wide range of studies of health and well-being via questionnaire and telephone interview studies, and recruitment was extended to their relatives (parents, siblings, adult children and spouses).

LifeLines Cohort Study

The LifeLines Cohort Study (57) is a multi-disciplinary prospective population-based cohort study from the Netherlands, 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. 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.

TwinsUK

For the UK, we use data from TwinsUK, the largest adult twin registry in the country with more than 12,000 respondents (58). The TwinsUK Study recruited white monozygotic (MZ) and dizygotic (DZ) twin pairs from the TwinsUK adult twin registry, a group designed to study the heritability and genetics of age-related diseases (www.twinsuk.ac.uk). These twins were recruited from the general population through national media campaigns in the UK and shown to be comparable to age-matched population singletons in terms of clinical phenotype and lifestyle characteristics.

STR

The Swedish Twin Registry (STR) was first established in the late 1950s to study the importance of smoking and alcohol consumption on cancer and cardiovascular diseases whilst controlling for genetic propensity to disease. Between 1998 and 2002, the STR conducted telephone interview screening of all twins born in 1958 or earlier regardless of gender composition or vital status of the pair. This effort is known as Screening Across the Lifespan Twin study (SALT). A subsample of SALT (≈10,000) was genotyped as part of the TwinGene project (41, 59) and we use the this information in the current study.

Fertility trends

Aggregate data to describe country specific fertility trends have been obtained from the Human Fertility Database (HFD, http://www.humanfertility.org/cgi-bin/main.php) and the Human Fertility Collection (HFC, http://www.fertilitydata.org/cgi-bin/index.php) if available.

Both data collections are joint projects of the Max Planck Institute for Demographic Research (MPIDR) in Rostock, Germany and the Vienna Institute of Demography (VID) in Vienna, Austria. The projects provide access to detailed and high-quality data on period and cohort fertility. The HFD is entirely based on official vital statistics. The HFC incorporates a variety of valuable fertility data from diverse, not necessarily official, data sources. All data are freely available after registration. We focused on fertility information for cohorts that was aggregated for individuals older than 45.

For the UK, official data on birth order have only been collected within marriage, and might be biased. We therefore relied 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%3A77–2631333. For Estonia, data on completed reproduction by age 45 was only available until the year 1962. For subsequent cohorts, however, there was an estimate of AFB available based on official statistics at the age of 40. For Australia, no official data on a time series of cohort specific AFB was available and the trends are based on the data used for analysis in this study.

Genotypes

We received genotype data from all cohorts, which we imputed according to the 1000 genome panel – except for TwinsUK from which we already received the imputed data.

Genetic-relatedness-matrix (GRM)

To estimate the genetic relatedness-matrix (GRM) the HapMap3 imputation panel was used as a reference set as it was optimized to capture common genetic variation in the human genome (60).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 5% after merging, lower minor allele frequency than 1% and which failed the Hardy-Weinberg equilibrium test for a

threshold of 10^{-6}. We merged the datasets subsequently applying QC again after merging each data set. 847,278 SNPs could be utilized to estimate the GRM between individuals. We used the software Plink (19) for the quality control and merging of the datasets and GCTA (18) to estimate the genetic relatedness matrix

The GRM A_{ik} is estimated for each pair of individuals j and k:

$$A_{jk} = \frac{1}{N}\sum_{i=1}^{N}\frac{(x_{ij} - 2p_i)(x_{ik} - 2p_i)}{2p_i(1 - p_i)}$$

where x_{ij} and x_{ik} is the number of copies of the reference allele for the i^{th} SNP of the j^{th} or k^{th} individual and p_i is the frequency of the reference allele and *N* the number of SNPs. If two individuals had a higher genetic relatedness than 0.05, one was excluded from the analyses to avoid bias due to environmental confounders amongst close relatives.

Phenotypes

Number of children ever born was available for all cohorts, but in ARIC and TwinsUK, however, only for women. NEB measures number of children a woman has given birth to or a man has fathered. It was either directly asked or we constructed it from questions on the date of birth of each child.

The measure is not perfectly harmonized across cohorts because some questionnaires explicitly exclude still-births (HRS, ARIC) while others remain undefined (TwinsUK asked in some waves: “How many children have you given birth to?”; EGCUT asked: “Do you have any biological children?”, and subsequently: “Fill in their names, gender and date of birth). In STR, LifeLines, QIMR as well as most of the waves of the TwinsUK, information on both the date of birth and death of the child was asked. In LifeLines and TwinsUK, we compared the live birth measure with number of children ever born and, as expected, given the diminishing mortality

rate in both datasets, less than 0.2% 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. We therefore expect no large bias due to the fact that in some countries still-births are excluded.

The questionnaires were furthermore heterogeneous in the maximum number of children that could be named. However, within each cohort, the maximum number of children has never been named more often than in 0.5 per cent of the cases and we do not anticipate that our results are influenced by this.

Information on AFB was available in all cohorts except for ARIC and the HRS. It was asked directly (in TwinsUK) or was constructed using the date of birth of the oldest child and the year of birth of the respondent.

Since fertility is strongly age dependent, we focus on women only with completed fertility history in reference to the phenotype. In general, the end of women's reproductive lifespan occurs around the age of 45 and for men at the age of 50, thus, we only included individuals beyond those ages in our analyses. Furthermore, in vitro fertilization (IVF) – often related to twinning and multiple births – can bias results if IVF compensates genetically based infertility. However, in our TwinsUK sample, only 60 women reported using IVF, who we did not include in the final analyses.

For all models, both phenotypes were standardized (Z-transformed) by cohort, year of birth and sex.

GREML Models

Common SNP heritability estimates (Model 1)

The genetic component underlying a trait is commonly quantified in terms of SNP-heritability (h^2_{SNP}) as the proportion of the additive genetic variance explained by common SNPs across the genome over the overall phenotypic variance (σ^2_P) of the trait:

$$h^2_{SNP} = \frac{\sigma^2_g}{\sigma^2_P}.$$

The phenotypic variance is the sum of additive genetic and environmental variance, i.e. $\sigma^2_P = \sigma^2_g + \sigma^2_e$ where σ^2_g is the additive genetic variance explained by all SNPs across the genome and σ^2_e is the residual variance.

The methods we applied have been detailed elsewhere (18, 20–22, 24). Briefly, we applied a linear mixed model

$$\mathbf{y} = \boldsymbol{X\beta} + \mathbf{g} + \mathbf{e}$$

where **y** is an Nx1 vector of dependent variables, N is the sample size, **β** is a vector for fixed effects of the overall mean (intercept), **g** is the Nx1 vector with each of its elements being the total genetic effect of all SNPs for an individual, and **e** is an Nx1 vector of residuals. The variance covariance matrix of the observed phenotypes is:

$$\mathbf{V} = \mathbf{A}\sigma^2_g + \mathbf{I}\sigma^2_e$$

We have g~ $N(0, \mathbf{A}\sigma^2_g)$ and e~ $N(0, \mathbf{I}\sigma^2_e)$, **A** is the genetic relationship matrix (GRM) estimated from SNPs and **I** is an identity matrix. The variance components are estimated using the restricted maximum likelihood (REML) approach.

Genes x Population (Model 2)

The genes x demographic birth cohort interaction model is a joint model estimating global genetic effects for the fertility traits, effective between and within samples (σ_g^2) and the averaged additional genetic effects within cohorts (σ_{gxp}^2):

$$\mathbf{V} = \mathbf{A}\sigma_g^2 + \mathbf{A_{gxp}}\sigma_{gxp}^2 + \mathbf{I}\sigma_e^2$$

where **A** is the genetic relatedness matrix and $\mathbf{A_{gxp}}$ is a matrix only with values for pairs of individuals within populations 1-7:

$$\mathbf{A} = \begin{bmatrix} A_{p1p1} & A_{p2p1} & A_{p3p1} & A_{p4p1} & A_{p5p1} & A_{p6p1} & A_{p7p1} \\ A_{p1p2} & A_{p2p2} & A_{p3p2} & A_{p4p2} & A_{p5p2} & A_{6p2p} & A_{p7p2} \\ A_{c1p3} & A_{p2p3} & A_{p3p3} & A_{p4p3} & A_{p5p3} & A_{p6p3} & A_{p7p3} \\ A_{c1p4} & A_{p2p4} & A_{p3p4} & A_{p4p4} & A_{p5p4} & A_{p6p4} & A_{p7p4} \\ A_{c1p5} & A_{p2p5} & A_{p3p5} & A_{p4p5} & A_{p5p5} & A_{p6p5} & A_{p7p5} \\ A_{p1p6} & A_{p2p6} & A_{p3p6} & A_{p4p6} & A_{p5p6} & A_{p6p6} & A_{p7p6} \\ A_{p1p7} & A_{p2p7} & A_{p3p7} & A_{p4p7} & A_{p5p7} & A_{p6p7} & A_{p7p7} \end{bmatrix}$$

$$\mathbf{A_{gxp}} = \begin{bmatrix} A_{p1p1} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & A_{p2p2} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & A_{p3p3} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & A_{p4p4} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & A_{p5p5} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & A_{p6p6} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & A_{p7p7} \end{bmatrix}$$

Genes x Demographic birth cohort (Model 3)

The genes x demographic birth cohort interaction model is a joint model estimating the universal genetic effects for the traits, effective between and within samples (σ_g^2) and the averaged additional genetic effects within defined birth cohorts (σ_{gxb}^2):

$$\mathbf{V} = \mathbf{A}\sigma_g^2 + \mathbf{A_{gxd}}\sigma_{gxd}^2 + \mathbf{I}\sigma_e^2$$

whereas **A** is the genetic relatedness matrix and $\mathbf{A_{gxd}}$ is a matrix only with values for pairs of individuals within the same demographic birth cohorts b1-2:

$$\mathbf{A_{gxb}} = \begin{bmatrix} A_{d1d1} & 0 \\ 0 & A_{d2d2} \end{bmatrix}$$

Genes x Population x Demographic birth cohorts (Model 4)

Finally, we applied a model including both interaction terms from above and an additional interaction term $\mathbf{A_{gxcxb}}$ which is 0 for all pairs of individuals living in different time periods or in different cohorts

$$\mathbf{V} = \mathbf{A}\sigma_g^2 + \mathbf{A_{gxp}}\sigma_{gxp}^2 + \mathbf{A}\sigma_{gxd}^2 + \mathbf{A_{gxpxd}}\sigma_{gxpxd}^2 + \mathbf{I}\sigma_e^2$$

whereas **A** is the genetic relatedness matrix, $\mathbf{A_{gxp}}$ is a matrix only with values for pairs of individuals within populations 1-7 (Model 2), $\mathbf{A_{gxd}}$ is a matrix only with values for pairs of individuals within the same demographic periods b1-2 (Model 3) and $\mathbf{A_{gxpxd}}$ is a matrix only with values for pairs of individuals with both the same demographic periods and the same populations.

Bivariate Model

For bivariate analyses (22, 24), we split the data into individuals born before and after the turning point in fertility postponement in AFB (see also Supplementary Information Table S5). Based on Model 2, we estimate a bivariate model with two GRMs:

$$\boldsymbol{V}\begin{bmatrix} \boldsymbol{f_b} \\ \boldsymbol{f_a} \end{bmatrix} = \begin{bmatrix} \boldsymbol{A_b}\sigma_{g\_b}^2 + \mathbf{A_{gxp\_b}}\sigma_{gxp\_b}^2 + \mathbf{I}\sigma_{e\_b}^2 & \boldsymbol{A_{a\_b}}\sigma_{g\_a\_b}^2 + \mathbf{A_{gxp\_a\_b}}\sigma_{gxp\_a\_b}^2 \\ \boldsymbol{A_{a\_b}}\sigma_{g\_a\_b}^2 + \mathbf{A_{gxp\_a\_b}}\sigma_{gxp\_a\_b}^2 & \boldsymbol{A_b}\sigma_{g\_a}^2 + \mathbf{A_{gxp\_a}}\sigma_{gxp\_a}^2 + \mathbf{I}\sigma_{e\_a}^2 \end{bmatrix}$$

where as $\boldsymbol{f_b}$ **and** $\boldsymbol{f_a}$ are vectors of length N_b *and* N_a of fertility phenotypes (NEB or AFB) of individuals born **b**efore or **a**fter the postponement transition started, with N being the respective sample size of the subsets. Variance components refer to those from Model 2, whereas the lower index _b indicates that they are estimated in the subset of individuals born before and index _a born after the start of the postponement transition. The index _b_a

denominates the covariance of variances components across subsets. The correlation of the genetic factors are estimated as:

$$r_{\sigma^2_{gxp\_a\_b}} = \sigma^2_{gxp\_a\_b} / \sqrt{\sigma^2_{gxp\_a} * \sigma^2_{gxp\_b}}$$

and

$$r_{\sigma^2_{g\_a\_b}} = \sigma^2_{g\_a\_b} / \sqrt{\sigma^2_{g\_a} * \sigma^2_{g\_b}}$$

Acknowledgement

The research leading to these results received funding from NWO (Dutch National Science Organization) (VIDI grant 452-10-012 to MCM) and the European Research Council via an ERC Consolidator Grant SOCIOGENOME (615603 awarded to MCM, see www.sociogenome.com). FCT additionally received funding from sciencestarter.de for the research visit to the University of Queensland. SHL was supported by the Australian National Health and Medical Research Council (1080157) and the Australian Research Council (DE130100614). The authors wish to thank the participants from all cohorts under study. We also wish to thank Patrick Präg, Mariana Bonnouvrier and Nicola Barban for useful comments on previous versions of this manuscript. For additional acknowledgments (participating cohorts), please see Supplementary Information text SI 1.

References

1. Kohler H-P, Rodgers JL, Miller WB, Skytthe A, Christensen K (2006) Bio-social determinants of fertility. *Int J Androl* 29(1):46–53.

2. Mills MC, Tropf FC (2015) The Biodemography of Fertility: A Review and Future Research Frontiers. *Kolner Z Soz Sozpsychol* 55(Special Issues Demography):397–424.

3. Milot E, et al. (2011) Evidence for evolution in response to natural selection in a contemporary human population. *Proc Natl Acad Sci* 108(41):17040–17045.

4. Byars SG, Ewbank D, Govindaraju DR, Stearns SC (2010) Natural selection in a contemporary human population. *Proc Natl Acad Sci* 107(suppl 1):1787–1792.

5. Kohler H-P, Rodgers JL, Christensen K (1999) Is fertility behavior in our genes? Findings from a Danish twin study. *Popul Dev Rev* 25(2):253–288.

6. Rodgers JL, Kohler HP, Kyvik KO, Christensen K (2001) Behavior genetic modeling of human fertility: Findings from a contemporary Danish twin study. *Demography* 38(1):29–42.

7. Tropf FC, Barban N, Mills MC, Snieder H, Mandemakers JJ (2015) Genetic influence on age at first birth of female twins born in the UK, 1919-68. *Popul Stud (NY)*:129–145.

8. Kirk KM, et al. (2001) Natural selection and quantitative genetics of life-history traits in Western women: a twin study. *Evolution (N Y)* 55(2):423–435.

9. Visscher PM, Brown MA, McCarthy MI, Yang J (2012) Five years of GWAS discovery. *Am J Hum Genet* 90(1):7–24.

10. Balbo N, Billari FC, Mills MC (2013) Fertility in advanced societies: A review of research. *Eur J Popul Eur Démographie* 29(1):1–38.

11. Mills MC, Rindfuss RR, McDonald P, te Velde E (2011) Why do people postpone parenthood? Reasons and social policy incentives. *Hum Reprod Update* 17(6):848–860.

12. Manolio TA, et al. (2009) Finding the missing heritability of complex diseases. *Nature* 461(7265):747–753.

13. Zaitlen N, et al. (2013) Using extended genealogy to estimate components of heritability for 23 quantitative and dichotomous traits. *PLoS Genet* 9(5):e1003520.

14. Polderman T, et al. (2015) Meta-analysis of the heritability of human traits based on fifty years of twin studies. *Nat Genet*:doi: 10.1038/ng.3285.

15. Visscher PM, Hill WG, Wray NR (2008) Heritability in the genomics era—concepts and misconceptions. *Nat Rev Genet* 9(4):255–266.

16. Fisher RA (1930) *The genetical theory of natural selection.* (Oxford University Press).

17. Hughes K, Burleson MH (2000) Evolutionary Causes of Genetic Variation in Fertility and other Fitness Traits. Genetic Influences on Human Fertility and Sexuality., eds Rodgers J, Row DC, Miller WB (Kluwer Academic Publisher, Massachusetts), pp 7–33. 1st Ed.

18. Yang J, Lee SH, Goddard ME, Visscher PM (2011) GCTA: a tool for genome-wide complex trait analysis. *Am J Hum Genet* 88(1):76–82.

19. Purcell S, et al. (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 81(3):559–575.

20. Visscher PM, Yang J, Goddard ME (2010) A commentary on "common SNPs explain a large proportion of the heritability for human height"by Yang et al.(2010). *Twin Res Hum Genet* 13(06):517–524.

21. Yang J, et al. (2010) Common SNPs explain a large proportion of the heritability for human height. *Nat Genet* 42(7):565–569.

22. Visscher PM, et al. (2014) Statistical power to detect genetic (co) variance of complex traits using SNP data in unrelated samples. *PLoS Genet* 10(4):e1004269.

23. Tropf F, Stulp G, Barban N, Visscher P (2015) Human Fertility, Molecular Genetics, and Natural Selection in Modern Societies. *PLoS One* 10(6):e0126821.

24. Lee SH, Yang J, Goddard ME, Visscher PM, Wray NR (2012) Estimation of pleiotropy between complex diseases using single-nucleotide polymorphism-derived genomic relationships and restricted maximum likelihood. *Bioinformatics* 28(19):2540–2542.

25. Kohler H-P, Billari FC, Ortega JA (2002) The emergence of lowest-low fertility in Europe during the 1990s. *Popul Dev Rev* 28(4):641–680.

26. de Kaa DJ Van (1987) Europe's second demographic transition. *Popul Bull* 42(1):1–59.

27. Lesthaeghe R (1995) The second demographic transition in Western countries: An interpretation. *Gend Fam Chang Ind Ctries*:17–62.

28. Schmidt L, Sobotka T, Bentzen JG, Andersen AN (2012) Demographic and medical consequences of the postponement of parenthood. *Hum Reprod Update* 18(1):29–43.

29. Sobotka T (2004) Is Lowest-Low Fertility in Europe Explained by the Postponement of Childbearing? *Popul Dev Rev* 30(2):195–220.

30. Katus K, Puur A, Põldma A (2004) Population-related Policies in Estonia in the 20th Century: Stages and Turning Points. *Finnish Yearb Popul Res* 40:73–103.

31. Zuk O, Hechter E (2012) The mystery of missing heritability: Genetic interactions create phantom heritability. *Proc Natl Acad Sci* 109(4):1193–1198.

32. Zhu Z, et al. (2015) Dominance genetic variation contributes little to the missing heritability for human complex traits. *Am J Hum Genet* 96(3):377–385.

33. Yang J, et al. (2015) Genetic variance estimation with imputed variants finds negligible missing heritability for human height and body mass index. *Nat Genet* 47(10):1114–1120.

34. Briley D, Harden K, Tucker-Drob E (2015) Genotype× Cohort Interaction on Completed Fertility and Age at First Birth. *Behav Genet* 45(1):71–83.

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

36. Stearns SC, Byars SG, Govindaraju DR, Ewbank D (2010) Measuring selection in contemporary human populations. *Nat Rev Genet* 11(9):611–622.

37. Montgomery GW, Zondervan KT, Nyholt DR (2014) The future for genetic studies in reproduction. *Mol Hum Reprod* 20(1):1–14.

38. Rietveld CA, et al. (2013) GWAS of 126,559 Individuals Identifies Genetic Variants Associated with Educational Attainment. *Science (80-)* 340(6139):1467–1471.

39. Rietveld CA, et al. (2014) Common genetic variants associated with cognitive performance identified using the proxy-phenotype method. *Proc Natl Acad Sci* 111(38):13790–13794.

40. Rietveld CA, et al. (2013) Molecular genetics and subjective well-being. *Proc Natl Acad Sci* 110(24):9692–9697.

41. Benjamin DJ, et al. (2012) The genetic architecture of economic and political preferences. *Proc Natl Acad Sci* 109(21):8026–8031.

42. Bongaarts J (1983) The proximate determinants of natural marital fertility. *Determ Fertil Dev Ctries A Summ Knowl* (15):85.

43. Mousseau TA, Roff DA (1987) Natural selection and the heritability of fitness components. *Heredity (Edinb)* 59(Pt 2):181–197.

44. Milot E, Pelletier F (2013) Human Evolution: New Playgrounds for Natural Selection. *Curr Biol* 23(10):R446–R448.

45. Courtiol A, Rickard I, Lummaa V (2013) The demographic transition influences variance in fitness and selection on height and BMI in rural Gambia. *Curr Biol* 23(10):884–889.

46. Kohler H-P, Rodgers JL, Christensen K (2002) Between nurture and nature: the shifting determinants of female fertility in Danish twin cohorts. *Biodemography Soc Biol* 49(3-4):218–248.

47. Bolund E, Hayward A, Pettay J, Lummaa V (2015) Effects of the demographic transition on the genetic variances and covariances of human life-history traits. *Evolution (N Y)* 69(3):747–755.

48. Stulp G, Barrett L, Tropf FC, Mills MC (2015) Does natural selection favour taller stature among the tallest people on earth? *Proc R Soc London B Biol Sci* 282(1806):20150211.

49. Miller WB (1992) Personality traits and developmental experiences as antecedents of childbearing motivation. *Demography* 29(2):265–285.

50. Pasta D, Miller WB (2000) A heritability study of childbearing motivation. *Genetic Influences on Human Fertility and Sexuality* (Springer US), pp 107–120.

51. Udry JR (1996) Biosocial models of low-fertility societies. *Popul Dev Rev* 22:325–336.

52. White A, Folsom A (1996) Community surveillance of coronary heart disease in the Atherosclerosis Risk in Communities (ARIC) Study: methods and initial two years' experience. *J Clin Epidemiol* 42(9):223–233.

53. Investigators A (1989) The atherosclerosis risk in community (aric) study: design and objectives. *Am J Epidemiol* 129(4):687–702.

54. Jackson R, et al. (1996) Differences between respondents and nonrespondents in a multicenter community-based study vary by gender and ethnicity. *J Clin Epidemiol* 49(12):1441–1446.

55. Juster F, Suzman R (1995) An overview of the Health and Retirement Study. *J Hum Resour*:S7–S56.

56. Leitsalu L, Haller T, Esko T (2014) Cohort profile: Estonian biobank of the Estonian Genome center, University of Tartu. *Int J ….*

57. Klijs B, Scholtens S, Mandemakers J (2015) Representativeness of the LifeLines Cohort Study. *PLoS One* 10(9):e0137203.

58. Moayyeri A, Hammond C, Spector TD (2013) Cohort profile: TwinsUK and healthy ageing twin study. *Int J Epidemiol* 42(1):76–85.

59. Lichtenstein P, et al. (2006) The Swedish Twin Registry in the third millennium: an update. *Twin Res Hum Genet* 9(06):875–882.

60. Consortium IH 3 (2010) Integrating common and rare genetic variation in diverse human populations. *Nature* 467(7311):52–58.

Tables

Table 1. Heritability estimates of the full GREML model and gene environment interaction models for number of children ever born (NEB) and age at first birth (AFB)

	Number of children ever born							
Model	1		2		3		4	
	Estimate (SE)	p-value	Estimate (SE)	p-value	Estimate (SE)	p-value	Estimate (SE)	p-value
σ_g^2/σ_P^2	0.038 (0.0097)	2.0 x10^{-5}	0.0070 (0.011)	0.26	0.00 (0.013)	0.50	0.00 (0.015)	0.50
$\sigma_{gxp}^2/\sigma_P^2$	--	--	0.15 (0.024)	6.0x10^{-12}	--	--	0.073 (0.036)	0.18
$\sigma_{gxd}^2/\sigma_P^2$	--	--	--	--	0.097 (0.017)	3.3x10^{-16}	0.064 (0.020)	5.9x10^{-4}
$\sigma_{gxpxd}^2/\sigma_P^2$	--	--	--	--	--	--	0.085 (0.045)	0.0030
N	31396							

	Age at first birth							
Model	1		2		3		4	
	Estimate (SE)	p-value	Estimate (SE)	p-value	Estimate (SE)	p-value	Estimate (SE)	p-value
σ_g^2/σ_P^2	0.053 (0.019)	0.0020	0.024 (0.022)	0.14	0.00 (0.030)	0.35	0.011 (0.028)	0.32
$\sigma_{gxp}^2/\sigma_P^2$	--	--	0.10 (0.039)	0.0037	--	--	0.00 (0.062)	0.50
$\sigma_{gxd}^2/\sigma_P^2$	--	--	--	--	0.084 (0.031)	4.6 x10^{-4}	0.00 (0.040)	0.50
$\sigma_{gxpxd}^2/\sigma_P^2$	--	--	--	--	--	--	0.18 (0.077)	0.0032
N	16109							

Note: σ_g^2/σ_P^2 = proportion of observed variance in the outcome associated with genetic variance across all environments, $\sigma_{gxp}^2/\sigma_P^2$ = proportion of observed variance in the outcomes associated with *additional* genetic variance within populations, $\sigma_{gxd}^2/\sigma_P^2$ = proportion of observed variance associated with *additional* genetic variance within demographic birth cohorts, $\sigma_{gxpxd}^2/\sigma_P^2$ = proportion of observed variance associated with *additional* genetic variance within populations and demographic birth cohorts, p-values are based on likelihood-ratio test comparing the full model with the model with one constraining the particular effect to be zero, all analyses include the first 20 Principal Components, outcomes are standardized for sex, birth year and country.

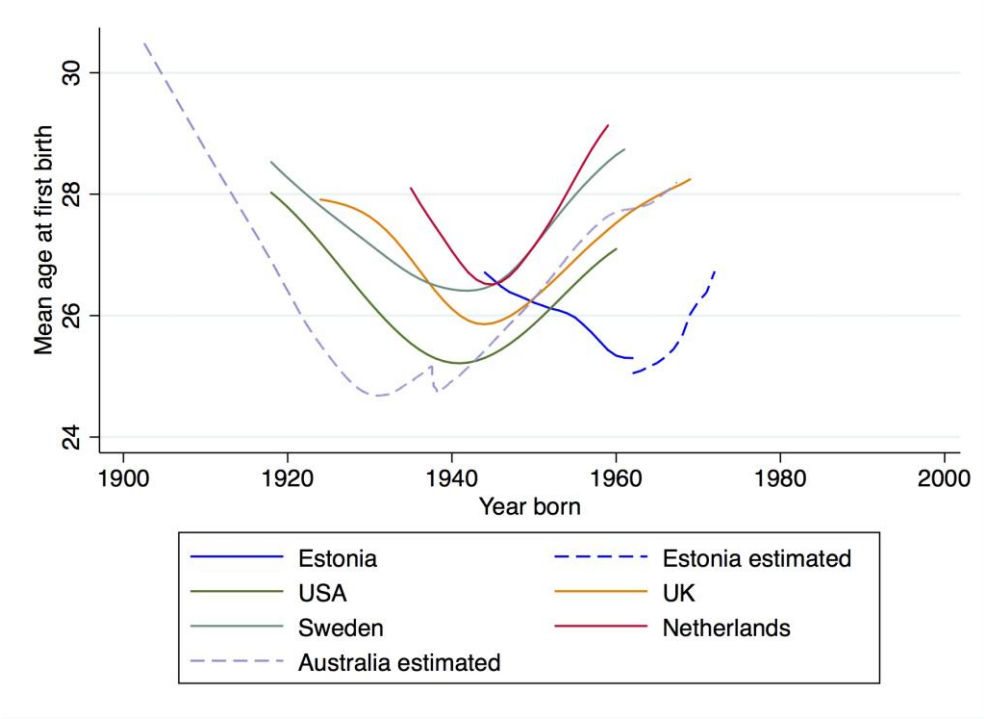
Table 2. Bivariate analysis of Model 2 to estimate genetic correlations for gxp (genes x population) or global g (gene) component before and after fertility postponement

	Number of children ever born						
	Before postponement		After postponement		r(G)		
	Estimate (SE)	p-value	Estimate (SE)	p-value	Estimate (SE)	p-value	N
$\sigma^2_{gxp}/\sigma^2_P$	0.15 (0.039)	$9.6x10^{-6}$	0.13 (0.048)	0.0010	1.00 (0.35)	$1.3x10^{-5}$	17,969
σ^2_g/σ^2_P	0.031 (0.018)	0.042	0.0017 (0.026)	0.50	-1.00 (8.04)	0.50	13,427
	Age at first birth						
	Before postponement		After postponement		r(G)		
	Estimate (SE)	p-value	Estimate (SE)	p-value	Estimate (SE)	p-value	N
$\sigma^2_{gxp}/\sigma^2_P$	0.099 (0.073)	0.083	0.13 (0.074)	0.070	0.11 (0.59)	0.27	8,049
σ^2_g/σ^2_P	0.023 (0.04)	0.20	0.011 (0.048)	0.50	1.00 (2.89)	0.50	8,060

Note: σ_g^2/σ_P^2 = proportion of observed variance in the outcome associated with genetic variance across all environments, $\sigma_{gxp}^2/\sigma_P^2$ = proportion of observed variance in the outcomes associated with *additional* genetic variance within populations, $\sigma_{gxd}^2/\sigma_P^2$ = proportion of observed variance associated with *additional* genetic variance within demographic birth cohorts, r(G) = genetic correlation, p-values are based on likelihood-ratio test comparing the full model with the model with one constraining the particular effect to be zero, all analyses include the first 20 Principal Components, outcomes are standardized for sex, birth year and country.

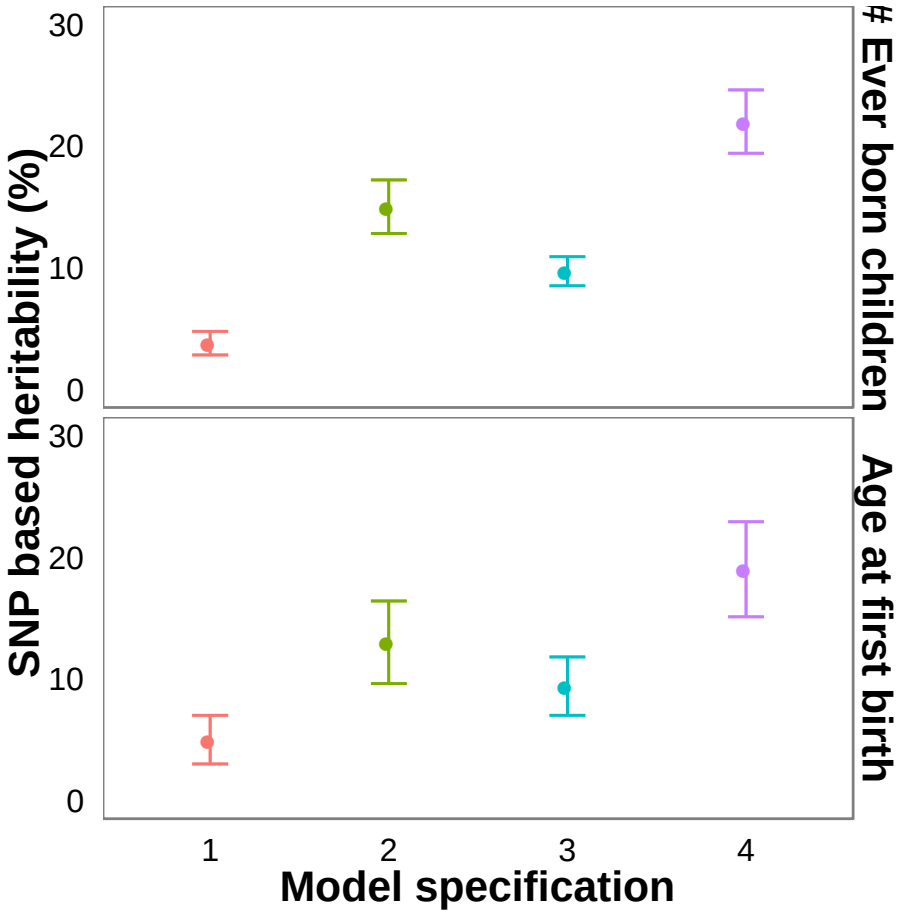
Figures

Figure 1. Trends in age at first birth in cohorts from the US, UK, Sweden, the Netherlands, Estonia and Australia (1903-1970).



Note: Trends in the mean age at first birth are moving averages based on aggregated data obtained from Human Fertility Database and the Human Fertility Collection (for details see Material and Methods). For Australia, no official data has been available and the trends have been estimated from the QIMR dataset. See Supplementary Fig. S2 for the birth cohort specific average in the QIMR data.

Figure 2. Bar Charts of the SNP-heritability estimates in number of children ever born (NEB) and age at first birth (AFB) for the different model specifications from Table 1



Note: SNP-heritability as the sum of genetic variance over the total variance in Model specification 1 = amongst all individuals, 2 = amongst individuals living within the same population, 3 = amongst individuals living within the same demographic birth cohort born either before or after fertility postponement, 4 = amongst individuals living in the same population and demographic birth cohort, dots = estimate, lines = estimate ± 1 SE, The corresponding table to Figure 2 an be found in Supporting Table S4.