

## University of Groningen

### Understanding childlessness

Verweij, Renske

**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:*

2019

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

*Citation for published version (APA):*

Verweij, R. (2019). *Understanding childlessness: Unravelling the link with genes and socio-environment*. [Thesis fully internal (DIV), University of Groningen]. Rijksuniversiteit 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.*

# CHAPTER

# 4

---

## Sexual dimorphism in the genetic influence on human childlessness\*

Renske M. Verweij, Melinda C. Mills, Felix C. Tropf, René Veenstra, Anastasia Nyman & Harold Snieder

*\*This chapter has been published as: Verweij, R. M., Mills, M. C., Tropf, F. C., Veenstra, R., Nyman, A., & Snieder, H. (2017). Sexual dimorphism in the genetic influence on human childlessness. European Journal of Human Genetics, (25), 1067–1074. doi:10.1038/ejhg.2017.105*

## ABSTRACT

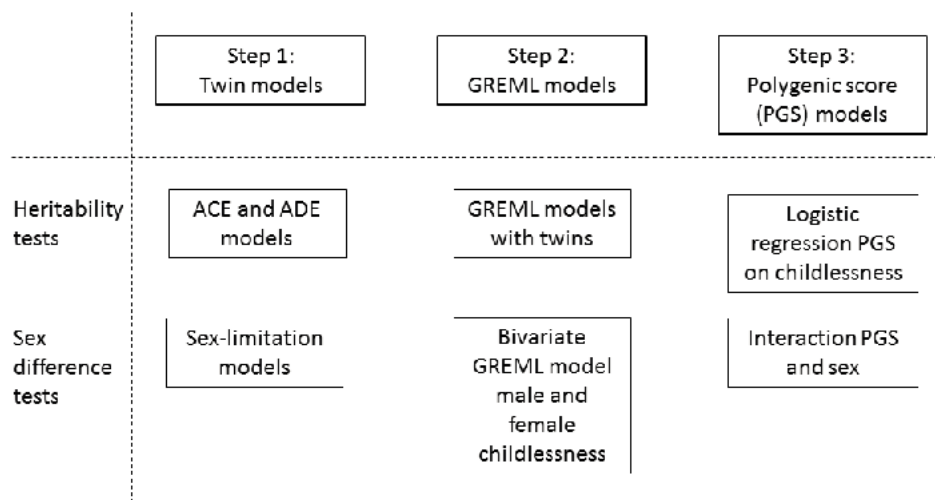
Previous research has found a genetic component of human reproduction and childlessness. Others have argued that the heritability of reproduction is counterintuitive due to a frequent misinterpretation that additive genetic variance in reproductive fitness should be close to zero. Yet it is plausible that different genetic loci operate in male and female fertility in the form of sexual dimorphism and that these genes are passed on to the next generation. This study examines the extent to which genetic factors influence childlessness and provides an empirical test of genetic sexual dimorphism. Data from the Swedish Twin Register (N= 9,942) is used to estimate a classical twin model, a genomic-relatedness-matrix restricted maximum likelihood (GREML) model on twins and estimates polygenic scores of age at first birth on childlessness. Results show that the variation in individual differences in childlessness is explained by genetic differences for 47% in the twin model and 59% for women and 56% for men using the GREML model. Using a polygenic score (PGS) of age at first birth (AFB), the odds of remaining childless are around 1.25 higher for individuals with 1 SD higher score on the AFB PGS, but only for women. We find that different sets of genes influence childlessness in men and in women. These findings provide insight into why people remain childless and give evidence of genetic sexual dimorphism.

## 4.1 INTRODUCTION

Over the last decades, human reproductive research has increasingly focused on biodemographic and genetic factors (Mills & Tropf, 2015). As child mortality diminished in contemporary societies, evolutionary researchers used childlessness and number of children as a proxy for reproductive fitness, which is the ability to pass on genes to subsequent generations. Additive genetic variance in fitness implies natural selection in populations, with the underlying assumption that alleles leading to higher reproductive success are passed on with a higher frequency in future generations (Courtiol, Tropf, & Mills, 2016). The erroneous misinterpretation of Fisher's Fundamental Theorem of Natural Selection that genetic variance in fitness should be close to zero has resulted in less attention to the study of genetics and reproduction (Fisher, 1930). Fisher argued that reproductive fitness is moderately heritable in humans, with a growing number of twin and family studies showing reproduction to be 25-50 percent heritable (Mills & Tropf, 2015). Previous research has found genetic influences on fecundity and reproductive desires (Mills & Tropf, 2015; Rodgers, Hughes, et al., 2001) with a recent GWAS isolating 12 genetic loci implicated in the timing and number of children (Barban et al., 2016).

Reasons for genetic effects on childlessness could be gene-environment interaction, non-additive genetic effects, or new mutations that restore any genetic variance lost to selection. Another hypothesis is that sexual dimorphism or in other words differences in secondary sex characteristics, operates since genes contributing to male childlessness are inherited via the female lineage and those for female childlessness via the male lineage (Hughes & Burlinson, 2000). There are likewise sex differences in biological makeup, processes and diseases implicated in infertility and behavior. For women, ovulatory problems, tubal damage, endometriosis, cervix cancer and polycystic ovary syndrome are prominent causes of infertility, with sperm defects and testicular cancer being central factors for men (Blundell, 2007). These diseases are partly heritable (Chen et al., 2011; Czene, Lichtenstein, & Hemminki, 2002). There is also a behavioural component to sexual dimorphism, since genes are implicated in different ways in relation to educational level and certain personality traits, including sociability, impulsivity and emotionality (Briley et al., 2017). These traits, which potentially have different effects on male and female fertility (Jokela, Kivimäki, Elovainio, & Keltikangas-Järvinen, 2009; Kravdal & Rindfuss, 2008; Nisén et al., 2013) have also been shown to have a heritable component in previous research (Nisén et al., 2013; G. E. Robinson, Fernald, & Clayton, 2008). Isolating the extent of sexual dimorphism in childlessness fosters a better understanding of why genetic variation in this trait still exists.

Data from the TwinGene project of the Swedish Twin Registry, which includes genotyped same sex and opposite sex twin pairs, is used to answer this question. This study extends previous research in three central ways. First, research on childlessness has been sparse in behavior genetics. Second, we also focus on men, who have been largely neglected in this area of research (Forste, 2002). Third, heritability (i.e., the proportion of variation in a trait within a specific population due to genetic variation), as well as sex differences, are estimated and contrasted using three different methods (see Figure 1 for an overview):



**Figure 1 |** Methods used to examine heritability and sex difference in heritability for childlessness

the classical twin method, the genomic-relatedness-matrix restricted maximum likelihood (GREML) method on twins and polygenic scores (PGS) from a recently published GWAS on timing and number of children (Barban et al., 2016) to assess the influence of SNPs (single-nucleotide polymorphisms) on childlessness for men and women at the molecular genetic level. SNPs are variations in a single nucleotide that occur at a specific position in the genome, where each variation is present to some degree within a population (Fagerness & Nyholt, 2008). The twin method enables us to compare results found in other countries and time periods. DZ twins share between 35% and 65% of their segregating genes (Visscher et al., 2006), which is assumed to be 50% in twin studies. The GREML method uses actual measured genetic similarity between twins, resulting in more precise estimates of heritability (see Supplementary Figure 2 for a comparison between methods). We use the polygenic scores to assess the extent to which a specific set of previously found SNPs influences childlessness differently in men and women.

## 4.2 MATERIALS AND METHODS

### 4.2.1 Data and zygosity

This study uses data from the Swedish Twin Registry (STR) which is compiled from the national birth registry (Magnusson et al., 2013). The STR contains multiple sub-studies of which we use the TwinGene project, where a subset of twins from the Screening Across the Lifespan Twin study (SALT) were genotyped. Between 1998 and 2002 data was collected from twins born between 1911 and 1958. Between 2004 and 2008 a subset of SALT participants were genotyped for the TwinGene project (Magnusson et al., 2013). We restrict

our analyses to TwinGene participants to ensure comparability of results between our different methodological techniques.

We include women aged 45 and older and men 50 and older, since less than 0.05% of children are born to mothers older than 45 (Billari et al., 2007) and no men in our sample had a first child after the age of 50. The TwinGene sample contains 10,909 individuals. After age restrictions, 9,942 individuals remained, with 4,534 men and 5,408 women coming from a total of 6,330 different families, 3,612 complete twin pairs and 2,718 individuals for whom there is only one twin from the pair in the sample.

Monozygotic (MZ) and dizygotic (DZ) twin pairs were distinguished using two different methods (Magnusson et al., 2013). If blood samples were available for both twins, genetic-based analyses were used to determine zygosity. This was not possible in two cases. In the TwinGene data, in some cases only one twin of a pair participated and was genotyped. Furthermore, in the case of presumed MZ pairs, also only one of each MZ pair was genotyped. In these two cases, zygosity was determined based on the responses to two questions. 'During childhood, were you and your twin partner as alike as "two peas in a pod" or not more alike than siblings in general?' and, 'How often did strangers have difficulty in distinguishing between you and your twin partner when you were children?'. If individuals responded 'as alike as two peas in a pod' and that strangers got confused 'almost always or always' or 'often', they were classified as MZ. This measure was tested by use of DNA markers and is accurate in 99% of the cases (Lichtenstein et al., 2002).

#### 4.2.2 Genotyping

The individuals were genotyped using the Illumina OmniExpress 700K chip, imputed to the 1000 genome imputation panel. After imputation we selected SNPs from the HapMap3 panel since this SNP set was optimized to capture common genetic variation (Altshuler et al., 2010), which is required for the GREML analysis. For quality control, SNPs with a minor allele frequency less than 1%, a higher missing rate than 0.03 and that failed the Hardy-Weinberg equilibrium for a threshold of  $10^{-6}$  were removed. The first 20 principal components are included as covariates to adjust for population stratification (Price et al., 2006).

#### 4.2.3 Measure of childlessness

Women are considered childless if they have no children who are living and no children who are dead (stillborn children are not counted as children who are dead). Men are considered childless if they have no children who are living. This results in a small discrepancy between the measurement of childlessness for men and women and for that reason we do robustness checks using only living children for men and women.

#### 4.2.4 Statistical Methods

To test heritability and sex differences of childlessness we used three methods: twin and GREML models and polygenic scores, illustrated in Figure 1.

*Twin method*

To quantify the genetic contribution to childlessness (a binary trait), we estimated a liability threshold model (M. C. Neale & Maes, 1992). This model assumes an underlying normal liability distribution that divides individuals into the two groups of childless versus not childless. Thresholds (z-values) for dividing these groups were estimated based on the proportion of childless individuals. The tetrachoric correlation of the liabilities in childlessness among MZ and DZ twin pairs was estimated using trait concordances (M. C. Neale & Maes, 1992). These correlations were then used to estimate the contribution of genetic and environmental factors in the same way covariances are used for continuous traits. In all twin models we control for birth year.

To test for genetic and environmental influences of childlessness, ACE, ADE, AE and CE models were fitted, which estimate the effect of additive genetic factors (A), non-additive/dominance genetic factors (D), shared (common) environmental factors (C) and individual (unique) environmental factors (E). The latter also contains measurement error. As shown in Supplementary Figure 1, A correlates 0.5 in DZ twin pairs and 1 in MZ twin pairs, D correlates 0.25 in DZ twin pairs and 1 in MZ twin pairs and C correlates 1 in both MZ and DZ twin pairs (Rijsdijk & Sham, 2002). When MZ correlations are more than twice the DZ correlations, an ADE model is estimated, which tests for dominant genetic effects, since D correlates perfectly for MZ twins but only 0.25 for DZ twins. Since they are confounded, C and D cannot be estimated simultaneously in univariate models. The A, C and D parameters were estimated using a model-fitting approach in which A, C and D factors were dropped in a stepwise fashion from the full model (ACE model or ADE model) and sub models were compared to the full model by hierarchical chi-square tests. The difference in the goodness-of-fit ( $-2 \log$  likelihood) between the sub- and full model is approximately chi-square distributed, with degrees of freedom equal to the difference in degrees of freedom. The model with the lowest Akaike's information criterion ( $AIC = X^2 - 2df$ ) reflects the optimal balance between goodness-of-fit and parsimony.

To examine quantitative and qualitative differences in the genetic and environmental etiology between males and females, sex limitation models were applied (Eley, 2005). Qualitative sex difference refers to whether different genes influence childlessness for males and females, tested by fitting a model in which the genetic correlation between opposite sex twin pairs was freely estimated and a model in which the genetic correlation is set at 0 (indicating independent genetic effects by sex) and by comparing these models to one in which the genetic correlation was fixed at 0.5 (indicating no sex-differences). A better fit of the model with the genetic correlation set to 0 thus indicates that different genes are implicated in male and female childlessness.

Quantitative sex difference refers to different proportions of additive genetic (A), shared environmental (C) and individual specific environmental (E) influence. We first ran a heterogeneity model in which the A, C and E parameters can differ between males and females followed by a homogeneity model where parameters for A, C and E are fixed as the same for the sexes. Differences between the goodness-of-fit of models were tested as described previously. For all twin models we used the OpenMx package in R (Boker et al., 2011).

*GREML method*

A second method used to estimate heritability is the genomic-relatedness-matrix restricted maximum likelihood (GREML) method on twins, which simultaneously considers the additive effect of all genotyped SNPs. The GREML method contains two steps. First, for each pair of individuals, the genetic similarity is estimated based on similarity in SNPs. Second, this genetic relatedness is used as the input as a random effect in a mixed linear model in which the genetic relatedness explains phenotypic similarity. This is done by a comparison of a matrix of pairwise genomic similarity to a matrix of pairwise phenotypic similarity (Yang et al., 2010). Since childlessness is a binary trait, the liability threshold model applies. The estimate of variance explained by SNPs on the observed scale is transformed to that on the underlying continuous scale (Sang Hong Lee, Wray, Goddard, & Visscher, 2011). We controlled for the first 20 principal components as well as birth year.

In this paper, we used a recently developed method that allows heritability to be estimated using both related and unrelated individuals (Zaitlen et al., 2013). We estimated narrow sense heritability ( $h^2$ ), commonly estimated in twin or family studies, and heritability based on genotyped SNPs ( $h^2_{snp}$ ). To estimate both  $h^2$  and  $h^2_{snp}$ , two covariance matrices were used: the identity-by-descent (IBD) and identity-by-state (IBS) matrices. The IBD matrix only includes individuals with relatedness above 0.05, for whom similarity of measured SNPs is an indicator of similarity over the whole genome. In the IBD matrix, genetic similarity for unrelated individuals (relatedness <0.05) is set to 0. The IBS matrix includes all individuals, but uses only information on unrelated individuals, because the information on related individuals is already captured by the IBD matrix. The IBS matrix thus captures only the genetic covariance for the SNPs in the genotyping array. We applied the joint model, which includes both the IBD and IBS matrices. The IBS matrix estimated  $h^2_{snp}$  and the IBD matrix estimates additional effects within families ( $h^2 - h^2_{snp}$ ), which together provide an estimate for narrow sense heritability ( $h^2$ ).

To examine whether the same or different genetic variants are implicated in male and female childlessness, bivariate GREML analysis were conducted with male childlessness considered as the first trait and female childlessness as the second trait, also used by Lee et al. for sex-differences in schizophrenia (S H Lee et al., 2012). The GCTA software (Yang, Lee, Goddard, & Visscher, 2011) was used for the GREML analysis.

*Polygenic scores*

The third method we used to assess the influence of genes on childlessness was creating the polygenic scores (PGS) for number of children ever born (NEB) and the age at which people have their first birth (AFB) and examine to what extent these PGS influence childlessness. The PGS is the sum of the risk alleles weighted by their effect size and is thus a summary measure of genetic variants that increase the risk for a trait (Wray et al., 2014). Different risk scores are created depending on p-value cutoffs, from using only genome wide significant SNPs (p-value of  $5 \times 10^{-8}$ ) to including all genotyped SNPs (p-value of 1). Polygenic scores are created with the PRSice tool in PLINK (Euesden, Lewis, & Reilly, 2015). An LD threshold of 0.1 and a distance threshold of 250kb are used, indicating that if two SNPs are included



in the PGS that have a correlation of 0.1 or greater, or a distance of 250kb or smaller, one of the two SNPs is removed. The original sample of the GWAS from which we create the PGS included the STR sample. For that reason, we used the GWAS results from the sample excluding STR and based our PGS on these results. We will run logistic regression models on childlessness with the standardized polygenic scores as independent variable controlling for year of birth and years of education. Only one individual from each twin pair is included in these analyses to meet the criteria of independent observations.

To assess sex differences in the effect of the polygenic risk score on childlessness fitted a logistic regression model including an interaction between the polygenic risk score and sex.

## 4.3 RESULTS

### 4.3.1 Background analysis

Around 12.6% of the women in the sample were childless, representative of childless women in Sweden, which has remained constant over the last decades at 13% (Statistics Sweden, 2011). Around 14.3% of the men in the sample were childless, which is lower than the overall rate in Sweden, which ranged between 17% and 20% in the period studied (Statistics Sweden, 2011). The correlation among MZ twins was in all cases higher than the correlation in DZ twins (see Table 1). This is an indicator that genetic factors play a role in childlessness. Among opposite sex twins, the tetrachoric correlation is  $-0.06$ . This is much lower than for same sex DZ twins, which was 0.17 for men and 0.28 for women (only the difference between opposite sex pairs and female DZ pairs is significant). This is a first indicator that the genetic or common environmental influence on childlessness differs between the sexes (Eley, 2005). We now discuss results from each method, summarized in Figure 2.

### 4.3.2 Results from the twin method

To estimate heritability in the twin model, univariate ACE models were estimated separately by sex. For males, ADE models were also estimated, since for male MZ twins, the tetrachoric correlation is more than twice as large as for male DZ twins, which is an indicator of dominant genetic effects. Both goodness-of-fit and parameter estimates for each model are listed in Table 2, with the best fitting models printed in bold. Comparing model 1 and 2, we see that for females, dropping C from the model does not significantly reduce model fit ( $p=0.796$ ) and when comparing model 1 and 3, dropping A resulted in a borderline significant reduction in model fit ( $p=0.072$ ). The best fitting model for females is thus model 2 – the AE model. The estimated heritability in this model was 0.48 (95% CI 0.33–0.62). For males, when contrasting model 5 with 6 and 9 with 10, we see that dropping C or D did not result in a significantly decreased model fit ( $p=1$  and 0.461 respectively). Comparing model 5 and 7 shows that dropping A resulted in a significant drop in model fit ( $p=0.017$ ), suggesting that the best fitting model for males is the AE (model 6). A heritability estimate of 0.46 (95% CI 0.30–0.61) indicates that almost half of the variance in childlessness is attributed to genetic

factors. For both sexes, there was no significant effect of shared environment, with the individual environment estimated slightly above 50%.

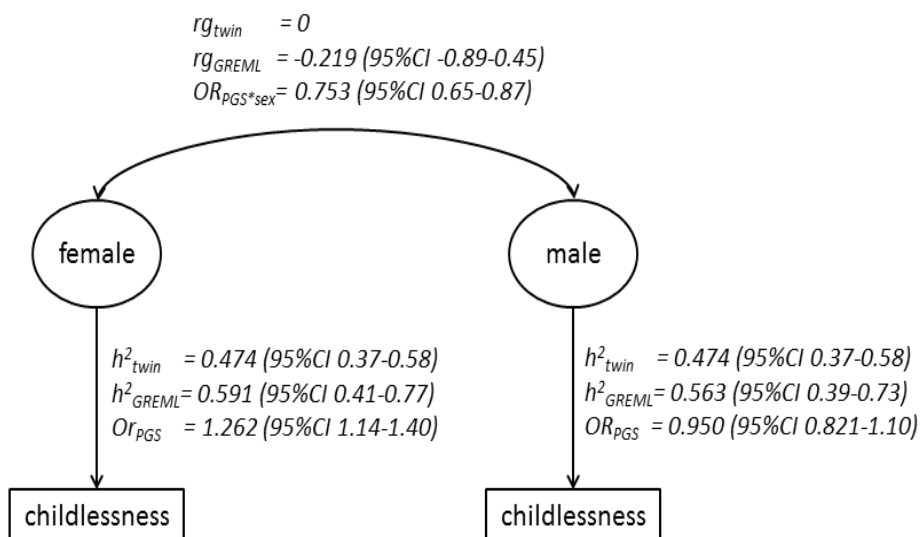
To examine whether there were different genetic influences on childlessness for males and females we fitted sex-limitation models. Goodness-of-fit statistics as well as parameter estimates are displayed in models 12 to 16 of Table 2. To examine qualitative sex differences, we tested whether the genetic correlation ( $r_g$ ) between men and women was different from the theoretical value of 0.5 or from 0. Since we did not find any shared environmental factors, we did not test whether the shared environmental correlation was different from the theoretical value of 1 and thus focused on testing the sex-limitations on our AE models. When the genetic correlation could be freely estimated in model 12, the estimate is 0.142. Model 13 in which the genetic correlation was set to 0.5 has a significantly lower model fit ( $p=0.023$ ). Model 14 in which the genetic correlation was set to 0 (which indicates that different genes play a role for males and females) has the best model fit, indicated by a value of  $-1.090$  for the AIC. For this reason, we adopt this model with the genetic correlation set to 0.

To test for quantitative sex differences, we examined whether the values for additive genetic influences (A) and individual environmental influences (E) could be set as equal between the sexes. This would indicate that the influence of the additive genetic and individual environment is equally important for men and women. Model 15, the homogeneous model, had a lower AIC value than the heterogeneous model 14 of  $-3.939$  and for that reason the values for A and E were set as equal between men and women.

To examine whether the estimated heritability from this model was significantly different from 0, in model 16 we tested if the A parameter could be dropped from the model. Dropping A significantly reduced model fit ( $p\text{-value}=0.000$ ), leaving the final best fitting model to be model 15. This was the model in which the genetic correlation between men and women was set to 0 without any effect of common environmental factors and equally high heritability estimates for men and women, with heritability estimated at 0.47 (95% CI 0.37-0.58) and individual environmental influences at 0.53 (95% CI 0.42-0.63) (see Table 2 and Figure 2). This indicates that there were no significant differences in the **extent** to which genes influence childlessness, but that there were qualitative genetic differences between male and female childlessness and that different genes influence childlessness in men compared to women.

#### 4.3.3 Results from the GREML method

In the next step, we examined heritability in twins using the GREML method (see Table 3). The estimated narrow sense heritability for the overall sample was 0.46 (95% CI 0.43-0.57). For females, the estimated narrow sense heritability was 0.59 (95% CI 0.41-0.77) and for males, 0.56 (CI 0.39-0.83) (see Figure 2). All estimates were significantly different from 0. The overall estimate is not the average of the male and female estimate, since male-female pairs were included in the overall analyses, which reduced heritability. Although the estimates from the GREML method are slightly higher than the twin model estimates, they do not significantly differ. The twin model estimates for both sexes is 0.474, which lies within the



**Figure 2** | Results for heritability and sex differences of childlessness from the twin, GREML and PGS (polygenic score) models. Twin estimates are from Table 2; model 15 (best fitting model) where the genetic correlation was set to 0 and the heritability estimate was set as equal between men and women. The model in which the genetic correlation was freely estimated was estimated at 0.14. GREML heritability estimates are taken from the model where heritability was estimated separately for men and women. Odds ratios come from Table 4; model 4, in which we use the genome wide genetic risk score (p-value of 1). The estimate for women is the main effect in this model and the estimate for men is the main effect\*the interaction for sex ( $1.262*0.753=0.950$ ).

95 confidence intervals of the GREML estimates for males (CI 0.394-0.732) and females (CI 0.413-0.769).

To further examine whether the same genes influence male and female childlessness, bivariate GREML models on childlessness were fitted to estimate the genetic correlation between childlessness by sex. The results are displayed in the bottom panel of Table 3 and in Figure 2. From the GREML analysis including twins, the genetic correlation between childlessness in males and females is  $-0.22$ , which is significantly different from 1 and not significantly different from 0. This indicates that, at least within this Swedish sample, a male and a female who have a higher genetic similarity do not have a higher similarity on childlessness. This shows that different genetic variants influence childlessness among males and females.

#### 4.3.4 Results from the PGS models

We then tested the effect of genes on childlessness by fitting logistic regression models on childlessness and testing the effect of the PGS of AFB and NEB. Table 4 and Figure 2 display the results for the AFB score. The models that use the PGS for NEB are not displayed since we did not find any significant results, which is not surprising since only 3 genetic loci

were significantly related to NEB. We display 4 models in Table 4. Model 1 includes the PGS including only genome wide significant SNPs, model 2 the PGS using all SNPs significant at the 0.05 level, model 3 all SNPs significant at the 0.5 level and model 4 all genotyped SNPs. For the PGS using only genome wide significant SNPs (p-value of  $5 \times 10^{-8}$ ) we find no significant effect on childlessness. For all other PGS's we find a significant effect with odds ratios of around 1.25. This indicates that the odds of remaining childless are about 1.25 times as high for individuals with a 1 standard deviation higher score on the AFB genetic risk score. Individuals with a greater risk of having a higher age at first birth are thus more often childless.

To test the sex differences in the effect of the polygenic risk score on childlessness, Table 4 and Figure 2 also display the results for the interaction between sex and the polygenic risk score. In all models except for the model that includes only genome wide significant SNPs, the interaction is significant and around 0.75. When looking at model 4, we see that the odds ratio for women is 1.262 and for men 0.950 ( $1.262 \times 0.753$ ). From this we can conclude that genes related to a higher age at first birth influence childlessness in women but not in men.

#### 4.3.5 Robustness checks

The measure of childlessness for men and women are not exactly the same. For women both living children as children who are dead are taken into account, while for men only children who are still living. Furthermore, only men over the age of 50 are included while women over the age of 45 are included. We fitted sex limitation models as well as the logistic regression models using the PGS on the measure on living children for men and women as well as on all men and women over 45 and over 50 to examine if this influences the results. Results are displayed in the supplementary material and in the supplementary tables, and show that neither the different age selection for men and women, nor the different measures of childlessness for men and women has a major impact on the results from our study. We furthermore show that the proportion of men who are considered childless because all their children died is relatively small.

## 4.4 DISCUSSION

The goal of this study was to examine sex-differences in the genetic influence on childlessness. We provide clear evidence that there are different genetic influences on childlessness for men and women. Although the level of the heritability of childlessness is approximately equal for both sexes, the actual genes that play a role vary. We infer this by applying classical twin modeling, the GREML method and a molecular genetic PGS approach. Future research should investigate which pathways genetic factors influence male and female childlessness. The question remains as to whether they are mainly physiological, behavioral, or whether gene-environment interactions work differently for men and women. For example, since women have a shorter reproductive window, the postponement of childbearing may have a larger impact on genetic factors influencing female childlessness.

We contrasted three different methods and compared their results in relation to male versus female childlessness. In the first classical twin method, we found that almost half of the variation (47%) in childlessness was due to genetic variation and that different genes influence male and female childlessness. We then applied the GREML method on twins. The main difference between the twin and GREML methods is that in the GREML method, genetic similarity between DZ twins is not assumed to be 50%, but measured on actual SNP similarity. Although the differences are not statistically significant, we find slightly higher heritability estimates of 59% with the GREML than the twin method, and also isolate that different genes influence male and female childlessness. Finally, using a PGS for AFB we found that genes previously found to be related to fertility timing are also related to childlessness for women, but not for men.

When comparing this study to previous twin studies on childlessness, we find comparable estimates of heritability in Finland (0.39 for women and 0.50 for men) (Nisén et al., 2013) and Denmark (for individuals born between 1880 and 1890 estimated at 0.45 for men and 0.70 for women and for individuals born between 1953 and 1964 estimates are 0.18 for men and 0.42 for women) (Kohler et al., 1999). One previous study on the STR found sex-limitations in genetic influences on the total number of children (Zietsch, Kuja-Halkola, Walum, & Verweij, 2014), which is in line with our findings. We extend this study, however, by looking at the different reproductive trait of childlessness instead of number of children or the age at first birth, use a broader birth cohort of Swedish twins born between 1911 and 1958 (instead of 1915-1930 in Zietsch and colleagues (2014)) and examine sex differences using three different methods.

Our findings somewhat contradict the recently published GWAS on human reproduction (age at first birth (AFB) and number of children ever born (NEB)), where only some sex-specific genetic effects in fertility were reported (section 5, SI) (Barban et al., 2016). In that study, out of the 12 independent loci isolated for human reproduction, two had a sex-specific effect. All signals found for AFB and two of the three signals for NEB had a consistent direction across the sexes. Using both LDscore bivariate regression and GREML bivariate analyses that study found a high genetic correlation among men and women for both traits. It is notable, however, that for AFB, the LD score regression results suggested that there were in fact sex-specific variants for AFB (i.e., the null hypothesis was rejected) and that genetic risk scores for NEB only significantly predicted childlessness in women and not in men (Barban et al. 2016, supplementary table 21). Another notable difference is that the GWAS examined continuous variables (i.e., AFB and NEB) and in this paper, we look at the binary outcome of childlessness. However, we do note that in our study we find much stronger sex differences, and more studies are necessary to confirm our conclusions and to clarify under which circumstances and for which fertility traits genes influence men and women differently.

We argue that different genes influence childlessness in males and females. A counterargument might be that differences in childlessness similarity in opposite sex twin pairs are not due to different genetic influences, but rather to different family socialization processes. However, we find no shared family influences in same sex twin pairs which

is in line with previous research that does not find that family characteristics such as sociodemographic background, family religiosity or socialization influence male and female fertility differently (Lytton & Romney, 1991). This makes it implausible that there are shared environmental family influences that make opposite sex siblings more dissimilar than same sex twin pairs. Furthermore, also our results from the PGS on unrelated individuals confirm our findings.

A shortcoming of this study is that we were unable to distinguish between voluntary (childfree) and involuntary childlessness, which might result in heterogeneity within the group of childless individuals. Genetic factors could influence the desire or predisposition to have children, biological fecundity or other pathways leading to childlessness. However, for the sake of examining whether genetic factors can be passed on to the next generation by sex differences, these findings are relevant regardless of our lack of distinction between (in)voluntary childlessness. A more general concern often raised with regard to twin studies, is the question of whether the trait of interest is the same amongst twins compared to the overall population. In this sample we find that the proportion of women who remain childless is equal to the overall proportion of childless women in Sweden. For men, the percentage in our sample is lower than the national percentage (Statistics Sweden, 2010). Previous research that examined this found no systematic differences between childlessness among twins and in the general population (Kohler, Knudsen, Skytthe, & Christensen, 2002). It is thus very likely that the lower percentage in our male sample is not attributed to the difference between twins and the general population, but rather differences in the measurement or response rates related to male reproduction. Another concern in twin studies is whether DZ twins share their environment to the same extent as MZ twins, referred to as the equal environment assumption (EEA). For several outcomes, this assumption has been tested by comparing the influence of perceived and actual zygosity, which gained plausible support (Plomin, Willerman, & Loehlin, 1976). Another study found that even though MZ twins share their environment to a higher extent than DZ twins, controlling for this rarely results in a significant reduction of the heritability estimate (Felson, 2014). Furthermore, previous research on unrelated individuals also found heritability of fertility traits (Tropf, Stulp, et al., 2015; Zaitlen et al., 2013). This indicates that estimates from the twin study might be an overestimate of the actual heritability, but that the overestimation is unlikely to be severe.

Finally, there are three concerns for potential inflation of heritability estimates in the GREML models we apply. First, ascertainment bias from the overrepresentation of cases in case-control studies cannot be corrected for if extended genealogical data is used (Zaitlen et al., 2013). However, given that in contrast to Zaitlen et al (2013) we only include pairs of twins in our study, this issue should not impact our results. Second, dominant genetic effects might bias narrow sense heritability estimates upwards in the GRM models (Zaitlen et al., 2013). Our twin models report no evidence for dominant genetic effects for childlessness – which is in line with the findings from recent reproductive (Mills & Tropf, 2015) and molecular genetics research (Barban et al., 2016). Third – and as discussed previously – shared environmental influences amongst siblings might influence fertility and correlate with genetic relatedness, inflating heritability estimates. Zaitlen et al (2013) find no evidence

for a bias due to shared environmental influences when introducing this approach to the literature for several traits including number of children – in line with findings from twin studies (Felson, 2014). Given the similarity of our findings from both the twin and GREML modeling approach, we are confident that our results are robust, also since twin models account for shared environmental influences and dominant genetic effects on childlessness.

This study has confirmed findings from previous research in showing a genetic influence on childlessness (Kohler et al., 1999), but uniquely identified genetic sexual dimorphism, which may be one of the explanations for previous findings on the moderate heritability of fertility.

## TABLES CHAPTER 4

**Table 1** | Concordance and tetrachoric correlations for childlessness in MZ and DZ twin pairs

Sample	N individuals	N complete pairs	Mean age	% childless	$C_{c/c}$	$D_{c/hc}$	Concordance (%)		Tetrachoric correlation (95% CI)
							CW	PW	
Female	MZ	1158	55.3	12	21	86	32.8	19.6	0.46 (0.28-0.62)
	DZ	2254	58.7	12.8	27	167	24.4	13.9	0.28 (0.12-0.42)
Male	MZ	1167	58.2	13.6	26	88	36.6	22.4	0.50 (0.33-0.64)
	DZ	1612	59.3	13.5	16	126	20.3	11.3	0.17 (-0.02-0.35)
Opposite sex	3751	1223	58.5	14.1	21	304	12.1	6.5	-0.06 (-0.18-0.09)
Total	9942	3612	58.2	13.4	111	501	22.1	12.4	0.22 (0.14-0.29)

$C_{c/c}$ = number of concordant childless twin pairs;  $D_{c/hc}$ = number of discordant childless twin pairs; CW= Case wise concordance; PW= Pairwise concordance. Tetrachoric correlation between childlessness in twin 1 and childlessness in twin 2.



**Table 2 |** Comparison of twin models and parameter estimates

Model		-2LL	DF	Comparison model	$\Delta$ -2LL	$\Delta$ DF	P-value	$\Delta$ AIC	$h^2_f$	$c^2_f$	$e^2_f$	$h^2_m$	$c^2_m$	$e^2_m$	$R_g$
1	Female	ACE	2539.67	3406					0.427	0.046	0.527				
<b>2</b>		<b>AE</b>	<b>2539.74</b>	<b>3407</b>	<b>1</b>	<b>0.07</b>	<b>0.796</b>	<b>-1.933</b>	<b>0.484</b>		<b>0.516</b>				
3		CE	2542.91	3407	1	3.24	0.072	1.244		0.343	0.657				
4		E	2573.63	3408	1	33.96	0.000	29.962			1.000				
5	Male	ACE	2172.13	2773							0.459	0.000	0.540		
<b>6</b>		<b>AE</b>	<b>2172.13</b>	<b>2774</b>	<b>5</b>	<b>0.00</b>	<b>1.000</b>	<b>-2.000</b>			<b>0.459</b>			<b>0.540</b>	
7		CE	2177.80	2774	5	5.68	0.017	3.678			0.328	0.671			
8		E	2201.54	2775	5	29.41	0.000	25.414			1.000				
9	Male	ADE	2171.58	2773							0.178	0.306	0.515		
<b>10</b>		<b>AE</b>	<b>2172.13</b>	<b>2774</b>	<b>9</b>	<b>0.54</b>	<b>1.000</b>	<b>-1.458</b>			<b>0.459</b>			<b>0.540</b>	
11		E	2171.79	2774	9	0.20	0.653	-1.799			1.000				
12	Qual.sex.diff	AE rg free	7747.17	9930					0.488		0.512	0.460		0.540	0.142
13		AE rg 0.5	7752.32	9931	12	5.15	0.023	3.150	0.410		0.590	0.444		0.556	
<b>14</b>		<b>AE rg 0</b>	<b>7748.08</b>	<b>9931</b>	<b>12</b>	<b>0.91</b>	<b>1.000</b>	<b>-1.090</b>	<b>0.460</b>		<b>0.540</b>	<b>0.488</b>		<b>0.512</b>	
									$h^2$	$c^2$	$e^2$				
<b>15</b>	Quant.sex. diff	AE rg0 M=F	7748.15	9933	14	0.07	2.000	-3.930	0.474	-	0.526				
16		E rg0 M=F	7811.38	9934	15	63.23	1.000	121.230	-	-	1.000				

All best fitting models are in bold. In all models birth year is controlled for. N female= 3412, N male=2779, N sex difference models = 9935, A=additive genetic; C=common environmental; E=individual environment; D=dominance genetic; -2LL= minus 2 log likelihood; DF=degrees of freedom;  $\Delta$ -2LL= difference in log likelihood, between the model and the comparison model;  $\Delta$ DF= difference between the degrees of freedom of the model and the comparison model; p-value for the chi-square test on  $\Delta$ -2LL with degrees of freedom from  $\Delta$ DF;  $\Delta$ AIC= difference in Akaike's information criterion between the model and the comparison model;  $h^2_f$ = heritability estimate female;  $c^2_f$ = estimate of shared environmental influence for female;  $e^2_f$ = estimate of individual environmental influence for female;  $h^2_m$ =heritability estimate men;  $c^2_m$ =estimate of shared environmental influence male;  $e^2_m$ =estimate of individual environmental influence male;  $R_g$ =estimate of genetic correlation

**Table 3** | GREML analysis on childlessness in the twin sample

Sample	Coeff	$h^2$	95% CI		N
Overall	$h^2$	0.455***	0.341	0.569	9942
Female	$h^2$	0.591***	0.413	0.769	5408
Male	$h^2$	0.563***	0.394	0.732	4534
<b>Rg</b>	$h^2-h^2_{snp}$	-0.219	-0.889	0.451	9942

\* $p < .05$ , \*\* $p < .01$ , \*\*\* $p < .001$ . Test if  $h^2/Rg$  is different from 0.  $h^2$  gives estimates for narrow sense heritability and Rg gives estimate of the genetic correlation.

**Table 4** | Results for the logistic regression models on childlessness using the polygenic risk scores for age at first birth using women and men over the age of 45 and 50 respectively

	OR	95% CI		p-value	
		Lower	Upper		
Model 1: AFB genes $p < 5 \times 10^{-8}$					
Intercept	1.05E+06	0.027	3.49E+13	0.118	
Years of education	0.993	0.970	1.016	0.984	
Birth year	0.992	0.983	1.001	0.542	
Sex (women=0, men =1)	1.134	0.979	1.313	0.077 .	
AFB PRS	0.999	0.899	1.110	0.094 .	
AFB PRS * Sex	1.034	0.894	1.197	0.651	
Model 2: AFB genes $p < 0.05$					
Intercept	3.68E+05	0.010	1.22E+13	0.148	
Years of education	0.992	0.969	1.015	0.487	
Birth year	0.992	0.984	1.001	0.099 .	
Sex (women=0, men =1)	1.148	0.991	1.332	0.066 .	
AFB PRS	1.216	1.095	1.351	0.000 ***	
AFB PRS * Sex	0.799	0.690	0.924	0.002 **	
Model 3: AFB genes $p < 0.5$					
Intercept	4.70E+05	0.012	1.56E+13	0.141	
Years of education	0.992	0.969	1.015	0.477	
Birth year	0.992	0.983	1.001	0.094 .	
Sex (women=0, men =1)	1.154	0.995	1.338	0.059 .	
AFB PRS	1.265	1.138	1.407	0.000 ***	
AFB PRS * Sex	0.753	0.651	0.871	0.000 ***	
Model 4: AFB genes $p \leq 1$					
Intercept	4.84E+05	0.013	1.60E+13	0.140	
Years of education	0.992	0.969	1.015	0.483	
Birth year	0.992	0.983	1.001	0.093 .	
Sex (women=0, men =1)	1.153	0.995	1.338	0.059 .	
AFB PRS	1.262	1.135	1.403	0.000 ***	
AFB PRS * Sex	0.753	0.651	0.872	0.000 ***	

\* $p < .05$ , \*\* $p < .01$ , \*\*\* $p < .001$ . N= 6614

## 4.5 SUPPLEMENTARY MATERIAL CHAPTER 4

### 4.5.1 Results from the robustness checks

The measures of childlessness for men and women are not exactly the same. Women are considered childless if they never had any children irrespective of whether these children are alive or have died, whereas men are considered childless if they don't have any living children. Furthermore, only men with age > 50y are included while for women this was age > 45y. For this reason, we fitted models using different age cutoffs (women >45y, men >50y; both > 45y and both >50y) and using different definitions of childlessness (the usual definition of childlessness, which is different for men and women vs. childlessness defined as having no living children).

Table S1 shows the tetrachoric correlations from our standard model (the first model: women >45y, men >50y, childless) and for the different age cutoffs and childlessness measures. No large differences are seen between estimates of the different models compared to our standard model.

Table S2 displays the sex limitation models for different age groups and childlessness measurements. Model 1 is our standard model as presented in the main text and is included for comparison. We can conclude that the results are very similar across all models. All heritability estimates from the final best fitting models are around 47%. The only slight discrepancy is that in Model 2 (both men and women > 45y), the best fitting model is the one in which the genetic correlation is estimated when testing for qualitative sex differences, in contrast to the others for which the model in which the genetic correlation was set to 0 was the best fitting. The estimated genetic correlation from this model is -0.09. However, this is still in line with our main findings that different genes influence childlessness in men and women.

Table S3 displays the Polygenic Risk Score (PGS) analysis using different definitions of age groups and childlessness measures. Again here we can see that all the methods show highly comparable results. In all models the influence of AFB genes is significant except for the first PGS that includes only genome-wide significant SNPs. For the other PGS the odds ratio is about 1.25 and there is a significant interaction with sex, which indicates that the effect of the SNPs for men is absent, very small or even negative and for women there is a positive significant effect.

From these robustness checks we can conclude that neither the different age selection for men and women, nor the different measures of childlessness for men and women has a major impact on the results from our study.

### 4.5.2 Expected probability of being childless due to the death of all your children

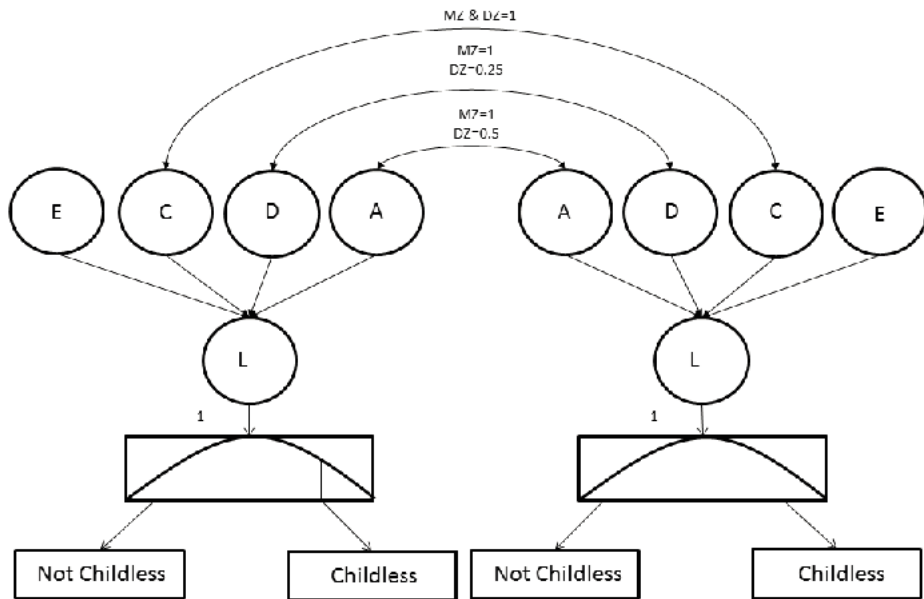
Table S4 displays the probability to be childless because all your children died at different ages. The probability to have 1, 2, 3 or 4 or more children is based on information from the human fertility database on cohort parity progression ratios among women born in Sweden in 1955 (Human Fertility Database, 2017c). The probability to have 1 child who died at different ages comes from the Swedish work and mortality database (Rostila, Saarela, &

Kawachi, 2012). This was on the level of the age of the child, which we transformed to the age of the parent by adding 28, which was the average age at birth in Sweden in the cohort of 1955 (Human Fertility Database, 2017b). The probability to have deceased 2 children is the multiplication of the probability to have one child who died ( $0.35 \times 0.35$  for two children at age 41,  $0.35 \times 0.35 \times 0.35$  for three children at age 41 etc.). The probability that all your children have died at a certain age is the sum of the probabilities to have a certain number of children times the probabilities that this number of children died.

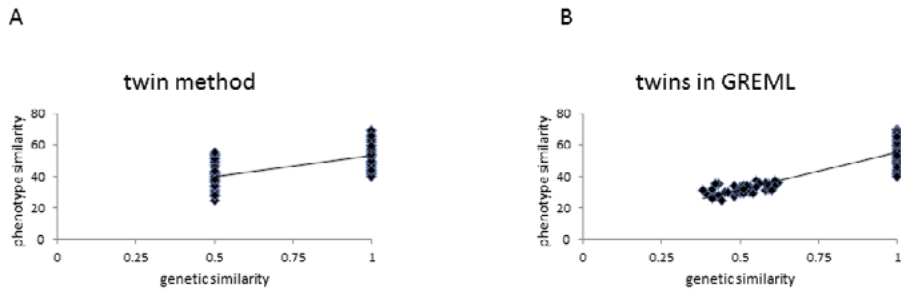
From this table we can see that at age 41, the probability to be childless because all your children died is 0.0057, at age 49 this probability is 0.0147, at age 57 this is 0.0181, at age 65 this is 0.0279 and at age 73 this is 0.0482.

These estimates are probably overestimates of the actual occurrences for a number of reasons. First of all, the group of 4+ children does not take larger number of children into account. If you would also include people with 5 or 6 children, the probability that all their children died is even lower for them, which would result in overall lower estimates. Secondly, the mean age at birth of 28 is used for women, because this is the data that is available. We also used this data to make predictions for men. However, because men often are a few years older than women when they have children, for men the proportion we give for all ages might only be achieved a few years later. Thirdly, we use data on fertility in the 1955 cohort, because this data is available. However, in older cohorts, who are primarily represented in our data, the fertility was higher, and for higher number of children the chance that all your children died is smaller. The estimate for number of children that died at different ages is probably appropriate to use for our sample, as the data represents births occurring between 1938 and 1980, which is also the time at which our sample, that was born between 1911 and 1958, was at childbearing age. We furthermore assume that deaths of children in families are independent of each other, i.e., the death of one child does not increase or decrease the probability that another child in the same family dies.

When comparing these estimates to what we find for women in our sample, we find that there are 15 women childless because all their children died, while there are 4707 women who have living children and 681 who never had any children. So from all women who ever had children, the probability to be childless because all their children died is 0.00317. This is indeed lower than our calculations in Table S4. In this sample it is not useful to compare estimates at different ages because the numbers are too small. In summary, we can conclude that even though our measurement of childlessness for men is not ideal, this measurement will not substantially deviate from actual childlessness in men and therefore only influence our results to a small extent.



**Supplementary Figure 1** | Path diagram for the liability threshold ACDE model. E, individual (unique) environmental component; C, Shared (common) environmental component; D, Dominance component; A, additive genetic component; L, liability. Please note that C and D cannot be estimated simultaneously.



**Supplementary Figure 2** | Graphical representation of the twin method (A) and the GREML method including twins (B). The figures display the relation between genetic similarity and phenotypic similarity in pairs of individuals (simulated data).

**Table S1 |** Tetrachorical correlations using different childlessness measures and age selection

	Tetrachoric correlation	95% CI	
		lower bound	upper bound
<b>Women 45 men 50 childless</b>			
MZ female	0.475	0.291	0.630
DZ female	0.266	0.109	0.413
MZ male	0.485	0.313	0.631
DZ male	0.166	-0.025	0.349
Opposite sex	-0.067	-0.204	0.072
<b>45 45 childless</b>			
MZ female	0.480	0.296	0.635
DZ female	0.279	0.123	0.425
MZ male	0.502	0.346	0.637
DZ male	0.161	-0.015	0.332
Opposite sex	-0.044	-0.172	0.087
<b>50 50 childless</b>			
MZ female	0.454	0.249	0.627
DZ female	0.284	0.119	0.437
MZ male	0.485	0.313	0.631
DZ male	0.166	-0.025	0.349
Opposite sex	-0.067	-0.204	0.072
<b>45 45 no living children</b>			
MZ female	0.466	0.281	0.622
DZ female	0.276	0.124	0.418
MZ male	0.518	0.364	0.649
DZ male	0.162	-0.016	0.332
Opposite sex	-0.031	-0.159	0.099
<b>50 50 no living children</b>			
MZ female	0.438	0.232	0.612
DZ female	0.264	0.101	0.415
MZ male	0.503	0.334	0.646
DZ male	0.166	-0.025	0.350
Opposite sex	-0.053	-0.190	0.085

**Table S2** | Robustness checks of the twin models: using different childlessness measures and age selection

Model 1: women 45+ men 50+ outcome: childless (as displayed in the main text)														
	-2LL	DF	Comparison	$\Delta$ -2LL	$\Delta$ DF	P-value	$\Delta$ AIC	$h^2_f$	$c^2_f$	$e^2_f$	$h^2_m$	$c^2_m$	$e^2_m$	$R_g$
Qualitative sex differences	7747.17	9930	AE rg free					0.488		0.512	0.460		0.540	0.142
	7752.32	9931	AE rg free	5.147	1	0.023	3.150	0.410		0.590	0.444		0.556	
	<b>7748.08</b>	<b>9931</b>	<b>AE rg free</b>	<b>0.907</b>	<b>1</b>	<b>0.341</b>	<b>-1.090</b>	<b>0.460</b>		<b>0.540</b>	<b>0.488</b>		<b>0.512</b>	
				$h^2$	$c^2$	$e^2$								
Quantitative sex differences	<b>7748.15</b>	<b>9933</b>	<b>AE rg 0 M=F</b>	<b>0.066</b>	<b>2</b>	<b>0.967</b>	<b>-3.930</b>	<b>0.474</b>		<b>0.526</b>				
	7811.38	9934	AE rg 0 M=F	63.235	1	0.000	121.230	-		1.000				
Model 2: women 45+ men 45+ outcome: childless														
Qualitative sex differences	<b>6148.22</b>	<b>7743</b>	<b>AE rg free</b>					0.498		0.502	0.474		0.526	<b>-0.090</b>
	6156.24	7744	AE rg free	8.028	1	0.005	8.028	0.418		0.582	0.438		0.562	
	6148.65	7744	AE rg free	0.430	1	0.512	0.429	0.474		0.526	0.498		0.502	
				$h^2$	$c^2$	$e^2$								
Quantitative sex differences	<b>6148.7</b>	<b>7746</b>	<b>AE rg 0 M=F</b>	<b>0.055</b>	<b>2</b>	<b>0.973</b>	<b>-3.944</b>	<b>0.485</b>		<b>0.515</b>				
	6220.63	7747	AE rg 0 M=F	71.925	1	0.000	96.925			1				
Model 3: Women 50+ men 50+ outcome: childless														
Qualitative sex differences	5417.9	6929	AE rg free					0.483		0.517	0.460		0.540	-0.143
	5422.92	6930	AE rg free	5.020	1	0.025	3.021	0.411		0.589	0.431		0.569	
	<b>5418.8</b>	<b>6930</b>	<b>AE rg free</b>	<b>0.905</b>	<b>1</b>	<b>0.341</b>	<b>-1.095</b>	<b>0.460</b>		<b>0.540</b>	<b>0.483</b>		<b>0.517</b>	
				$h^2$	$c^2$	$e^2$								
Quantitative sex differences	<b>5418.85</b>	<b>6932</b>	<b>AE rg 0 M=F</b>	<b>0.041</b>	<b>2</b>	<b>0.980</b>	<b>-3.958</b>	<b>0.471</b>		<b>0.529</b>				
	5476.42	6933	AE rg 0 M=F	57.571	1	0.000	55.570			1.000				

Table S2 | Continued

Model 4: Women 45+ men 45+ outcome: no living children													
Qualitative sex differences	AE rg free	6198.2	7745	AE rg free	9,199	1	0.002	7,199	0.486	0.514	0.488	0.512	-0.063
	AE rg 0.5	6207.4	7746	AE rg free	0.212	1	0.645	-1.788	0.431	0.569	0.420	0.580	
	<b>AE rg 0</b>	<b>6198.41</b>	<b>7746</b>	<b>AE rg free</b>	<b>0.212</b>	<b>1</b>	<b>0.645</b>	<b>-1.788</b>	<b>0.488</b>	<b>0.512</b>	<b>0.486</b>	<b>0.514</b>	<b>0.514</b>
									$h^2$	$c^2$		$e^2$	
Quantitative sex differences	<b>AE rg 0 M=F</b>	<b>6198.41</b>	<b>7748</b>	<b>AE rg 0</b>	<b>0.000</b>	<b>2</b>	1.000	-5.788	0.487	0.513		0.513	
	E rg 0 M=F	6272.01	7749	AE rg 0 M=F	73.602	1	0.000	71.602	1.000				
Model 5: Women 50+ men 50+ outcome: no living children													
Qualitative sex differences	AE rg free	5465.58	6930	AE rg free	5,832	1	0.016	3,832	0.461	0.539	0.476	0.524	0.113
	AE rg 0.5	5471.41	6931	AE rg free	0.561	1	0.454	-1.439	0.428	0.572	0.400	0.600	
	<b>AE rg 0</b>	<b>5466.14</b>	<b>6931</b>	<b>AE rg free</b>	<b>0.561</b>	<b>1</b>	<b>0.454</b>	<b>-1.439</b>	<b>0.476</b>	<b>0.524</b>	<b>0.461</b>	<b>0.539</b>	<b>0.539</b>
									$h^2$	$c^2$		$e^2$	
Quantitative sex differences	<b>AE rg 0 M=F</b>	<b>5466.15</b>	<b>6933</b>	<b>AE rg 0</b>	<b>0.018</b>	<b>2</b>	<b>0.991</b>	<b>-3.982</b>	<b>0.469</b>	0.531		0.531	
	E rg 0 M=F	5524.04	6934	AE rg 0 M=F	57.886	1	0.000	55.885	1.000			1.000	

All best fitting models are in bold. In all models birth year is controlled for



**Table S3** | Robustness checks for the PGS models: results for the logistic regression models on childlessness using the polygenic risk scores for age at first birth using different childlessness measures and age selection

	Model AFB genes $p < 5 \times 10^{-8}$		Model AFB genes $p < 0.05$		Model AFB genes $p < 0.5$		Model AFB genes $p \leq 1$	
	OR	95% CI	OR	95% CI	OR	95% CI	OR	95% CI
<b>Model 1: women 45+ men 50+ outcome: childless (as displayed in the main text)</b>								
Intercept	1.05E+06	(0.027-3.49E+13)	3.68E+05	(0.010-1.22E+13)	4.70E+05	(0.012-1.56E+13)	4.84E+05	(0.013-1.60E+13)
years of education	0.993	(0.970-1.016)	0.992	(0.969-1.015)	0.992	(0.969-1.015)	0.992	(0.969-1.015)
birth year	0.992	(0.983-1.001)	0.992	(0.984-1.001)	0.992	(0.983-1.001)	0.992	(0.983-1.001)
sex (women=0, men =1)	1.134	(0.979-1.313)	1.148	(0.991-1.332)	1.154	(0.995-1.338)	1.153	(0.995-1.338)
AFB genes	0.999	(0.899-1.110)	1.216	(1.095-1.351)***	1.265	(1.138-1.407)***	1.262	(1.135-1.403)***
AFB PRS 1*sex	1.034	(0.894-1.197)	0.799	(0.690-0.924)**	0.753	(0.651-0.871)***	0.753	(0.651-0.872)***
N	6310							
<b>Model 2: women 45+ men 45+ outcome: childless</b>								
Intercept	0.063	(0.000-7.89E+05)	0.023	(0.000-2.90E+05)	0.024	(0.000-3.03E+05)	0.024	(0.000-3.05E+05)
years of education	0.995	(0.973-1.017)	0.994	(0.972-1.017)	0.994	(0.971-1.016)	0.994	(0.971-1.016)
birth year	1	(0.992-1.009)	1.001	(0.993-1.010)	1.001	(0.993-1.010)	1.001	(0.993-1.010)
sex (women=0, men =1)	1.227	(1.065-1.414)**	1.246	(1.081-1.437)**	1.25	(1.085-1.443)**	1.25	(1.084-1.443)**
AFB genes	0.994	(0.894-1.106)	1.221	(1.099-1.359)***	1.264	(1.136-1.407)***	1.26	(1.133-1.403)***
AFB PRS 1*sex	1.103	(0.957-1.271)	0.802	(0.696-0.924)**	0.77	(0.668-0.888)***	0.772	(0.670-0.890)***
N	6544							
<b>Model 3: women 50+ men 50+ outcome: childless</b>								
Intercept	1.47E+04	(0.000-2.29E+12)	3.56E+03	(1.80E-05-5.70E+11)	3.73E+03	(0.000-6.00E+11)	3.70E+03	(0.000-5.94E+11)
years of education	0.993	(0.970-1.017)	0.992	(0.969-1.016)	0.992	(0.968-1.016)	0.992	(0.969-1.016)
birth year	0.994	(0.984-1.004)	0.995	(0.985-1.005)	0.995	(0.985-1.005)	0.995	(0.985-1.005)
sex (women=0, men =1)	1.113	(0.958-1.294)	1.126	(0.968-1.311)	1.132	(0.973-1.318)	1.131	(0.972-1.317)
AFB genes	1	(0.894-1.117)	1.215	(1.086-1.359)***	1.268	(1.133-1.420)***	1.261	(1.127-1.413)***

Table S3 | Continued

	Model AFB genes $p < 5 \times 10^{-8}$		Model AFB genes $p < 0.05$		Model AFB genes $p < 0.5$		Model AFB genes $p \leq 1$	
	OR	95% CI	OR	95% CI	OR	95% CI	OR	95% CI
AFB PRS 1*sex	1.036	(0.891-1.204)	0.8	(0.688-0.930)**	0.76	(0.653-0.883)***	0.762	(0.655-0.886)***
N	5858							
Model 4: women 45+ men 50+ outcome: no living children								
Intercept	7.50E+05	(0.022-2.23E+13)	2.78E+05	(0.008-8.26E+12)	9.57E+12	(0.009-9.57E+12)	3.29E+05	(0.010-9.78E+12)
years of education	0.993	(0.971-1.016)	0.992	(0.970-1.015)	0.992	(0.969-1.015)	0.992	(0.969-1.015)
birth year	0.992	(0.983-1.001)	0.993	(0.984-1.002)	0.993	(0.984-1.001)	0.993	(0.984-1.001)
sex (women=0, men =1)	1.108	(0.957-1.282)	1.12	(0.967-1.297)	1.125	(0.972-1.304)	1.126	(0.972-1.304)
AFB genes	0.998	(0.899-1.108)	1.195	(1.078-1.327)***	1.251	(1.127-1.390)***	1.249	(1.125-1.388)***
AFB PRS 1*sex	1.034	(0.894-1.195)	0.816	(0.706-0.943)**	0.765	(0.662-0.884)***	0.764	(0.661-0.883)***
N	6313							

\* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ .

**Table S4** | Probability to be childless due to the death of all your children, at different ages

Age	41	49	57	65	73
Probability to have:					
1 child	0.1491	0.1491	0.1491	0.1491	0.1491
2 children	0.4106	0.4106	0.4106	0.4106	0.4106
3 children	0.2230	0.2230	0.2230	0.2230	0.2230
4+ children	0.0893	0.0893	0.0893	0.0893	0.0893
Probability to have:					
1 child who died	0.0350	0.0800	0.0950	0.1340	0.2000
2 children who died	0.0012	0.0064	0.0090	0.0180	0.0400
3 children who died	0.0000	0.0005	0.0009	0.0024	0.0080
4 children who died	0.0000	0.0000	0.0001	0.0003	0.0016
probability that:					
all the children you had died	0.0057	0.0147	0.0181	0.0279	0.0482



