Social science genetics and fertility

Tropf, Felix

1. Introduction to social science genetics & overview of the studies[1]

Fertility research on human populations has been largely dominated by social science or environmental explanations of fertility behaviour and outcomes (Balbo, Billari, & Mills, 2013). Yet a growing body of research over the last few decades has demonstrated the relevance of including biological and genetic factors into our understanding of fertility outcomes (Foster, 2000; Kohler, Rodgers, & Christensen, 1999; Kohler, Rodgers, Miller, Skytthe, & Christensen, 2006; Kohler, Rodgers, & Christensen, 2002; Kohler & Rodgers, 1999, 2003; W B Miller, Bard, Pasta, & Rodgers, 2010; Rodgers et al., 2008; Tropf, Barban, Mills, Snieder, & Mandemakers, 2015; Tropf, Stulp, et al., 2015; Udry, 1996). Recent advances in biology, molecular genetics, medical sciences, reproductive medicine and evolutionary anthropology have likewise increased the relevance of adopting an interdisciplinary approach to study fertility.

The aim of this dissertation is to challenge and extend existing knowledge on genetic and environmental influences on fertility, and to provide an interdisciplinary approach across the fields of genetics and social sciences for fertility researchers. The main body of this book consists of four empirical studies, which quantify genetic effects on age at first birth and number of children based on twin data (chapters 2 & 5) and molecular genetic data (chapters 3 & 4), as well as the genetic correlation across both outcomes (chapter 3). We investigated patterns of gene-environment interaction across birth cohorts (chapters 2 & 4) as well as the common base of genetic effects on fertility outcomes across populations (chapter 3). Finally, we used genetically informed statistical models to approach the causal effect of education on the age at first birth as well as its contribution to fertility postponement during the 20th century in the United Kingdom (chapter 5).

1 This chapter is based on Mills, M. C. & F. C. Tropf (2016). The biodemography of fertility: A review and future research frontiers. In Social Demography Forschung an der Schnittstelle von Soziologie und Demografie (pp.397-424). Springer Fachmedien Wiesbaden.

This introduction has four goals: first, it provides some background on the history of research into the intersection between social science and biology, in order to make the book accessible to a readership from the social sciences. This review focuses on the contribution from biodemographic research. Biodemography is a new branch of science that integrates biology and demography, focussing on the complementary biological and demographic determinants of and interactions between the birth and death processes insofar as they relate to populations, and often to humans in particular (Carey & Vaupel, 2005). We first reflect on past challenges that slowed the application of a biodemographic approach to fertility research, and briefly describe the emergence of the field and introduce the basic concepts and terminologies. Second, we provide a review of previous examinations to take stock of what we have learned until now. Our work draws primarily on research carried out within demography and sociology, but with attention to more recent work particularly in the areas of evolutionary anthropology, and behavioural and molecular genetics. We will localize our work in this body of literature to identify the contributions of this dissertation in the larger framework of the scientific field. Third, we provide short summaries of the studies conducted in the chapters of this volume. Finally, we discuss persisting issues of our field of research and propose future frontiers for promising research in social science genetics and fertility in general.

We have opted to organize the scientific history of the biodemography of fertility around several themes. We first review the main reasons for the lack of attention – and in some cases outright resistance – to adopting a integrative approach to fertility research. We then discuss the early foundations of combining a biological and behavioural approach to the study of fertility via the use of intermediate fertility variables. Next, we turn to a review of research that has adopted a behavioural genetic approach to determine whether fertility has a genetic component, often using family and twin study designs. This is followed by more recent research in the area of molecular genetics, which gradually shifts from identifying whether there is a genetic component to fertility to isolating where it is located on the genome. Anthropological and evolutionary approaches are then touched upon. The chapter concludes with a broader discussion that reflects upon

what we have learned up until now, as well as potential future frontiers in this area of research.

1.1 A biodemographic approach to fertility

1.1.1 Defining fertility

Before embarking upon this volume, it is useful to first define the terminology related to 'fertility', as it often differs across disciplines. It is essential to note that the broader term fertility has different meanings in demography, evolutionary biology and reproductive medicine (Mills, Rindfuss, McDonald, & te Velde, 2011). In demography and related social sciences, 'fertility' refers to performance and the bearing and timing of live births. The focus is often on the two interrelated terms of the *tempo* (or timing) of childbearing and the *quantum* or actual number of children that women have during a certain period. Throughout this volume, We mainly refer to two of the most central indicators studied in this field until now, which is the tempo measure of age at first birth (AFB) and the quantum measure of number of children ever born (NEB).

In demography and sociology, quantum is often referred to as the number of children (e. g. Kohler et al., 1999) whereas in biological research the same outcome is referred to as life-time reproductive success (Stearns, Byars, Govindaraju, & Ewbank, 2010) or the number of offspring (Zietsch, Kuja-Halkola, Walum, & Verweij, 2014). In evolutionary research, fertility quantum may represent a proxy for 'fitness,'[2] which is a function of the number of children of a subject in relation to the number of children of peers of the same birth cohort (e. g. Kirk et al., 2001; Stearns et al., 2010). This in turn is used to measure how far the fertility quantum leads to relatively higher chances to successfully transmit genes to the next generation.

[2] We are aware that this is a simplification of the term fitness and that more nuanced measures differentiate between absolute fitness (ratio between number of individuals with genotype after selection to those before selection) and relative fitness (average number of surviving progeny of a particular genotype compared to average of surviving progeny of competing genotypes after a single generation). That is, fitness is defined as a propensity or probability of the expected average number of offspring in a particular group or class of individuals and not only the actual number of offspring produced. Additionally, long term measures such as number of grandchildren give a better assessment of reproductive success across generations.

Ideally, fitness of individuals is measured across several generations, however, it has been argued that due to improvements in hygiene and the reduction in prenatal, infant and child mortality in industrialized societies, NEB is a valid proxy for lifetime reproductive success indicating biological fitness—while much more feasible for empirical investigations (Stearns et al., 2010). Complementarily Zietsch and colleagues (2014) found a perfect genetic correlation between number of children and number of grandchildren amongst Swedish Twins, validating this argument. Another distinction between disciplines is the terminology related to fertility versus fecundity. In reproductive medicine, fertility denotes the ability of couples, women or men, to conceive and have children given unprotected intercourse. In demography and sociology, fertility refers to the number of children, while fecundity defines the ability to conceive.

1.1.2 Why is a socio-genetic approach to fertility less prevalent?

Although there has been some recognition of the biology underlying fertility, sociologists and demographers have been reticent in adopting and integrating findings and approaches from behavioural and molecular genetics, neuro-endocrinology and cross-species life history analysis (Wachter & Bulatao, 2003). Although not exhaustive, we outline some central reasons for this lack of enthusiasm.

A first historical reason, also noted by others (Kohler et al., 1999; Rodgers, Kohler, Kyvik, & Christensen, 2001), is the putative implication of Fisher's (1930) Fundamental Theorem of Natural Selection (FTNS). Fisher's theory states that fertility is a fitness trait, which theoretically entails that heritability should be (close to) zero. As we demonstrate shortly, however, a series of studies produced evidence that this is not the case.

A second prominent reason for the paucity of research that adopts a behavioural genetic approach in the social sciences in general – but particularly in relation to fertility – is the dark history related to eugenic policies that emerged in the 1880s and subsequent extreme historical atrocities. Eugenics focused on 'improving' humanity via supposedly scientific methods that proposed selective breeding. As Levine and Bashford (Furstenberg Jr., Levine, & Brooks-Gunn, 1990) describe, the aim of the eugenics movement was "to affect reproductive practice

through the application of theories of heredity." The aim was to prevent life (sterilization, contraception, abortion), encourage 'fitter' life (training, rearing of children, public health) and promote pro-natalist goals, but also at its most extreme, to end life (so-called euthanasia of the disabled) (Levine & Bashford, 2010). As a result of these policies, hundreds of thousands of people were segregated and sterilized, or lost their lives. This perspective has been widely, and rightly, condemned. It is essential to note that the type of research described in this volume and within the mainstream of contemporary peer-reviewed research in behavioural and molecular genetics has no eugenic goals or ties. Considering this grave history and link of eugenics and fertility, it remains important to explicitly acknowledge this point with the goal to prevent similar abuses to ever occur in the future.

Third, as noted in a *Population Studies* article by Thoday et al. (Thoday et al., 1970), social scientists often ignore biology and genetics (and vice versa) due to a lack of understanding and training in their concepts and methods, and to the virtual absence of cooperation between disciplines, which arguably still holds today almost 50 years later. As touched upon shortly, the growth of candidate gene studies in the social sciences came at around the same time that they were shown to be an incorrect method that produced false positives in genetics (Ioannidis, Ntzani, Trikalinos, & Contopoulos-Ioannidis, 2001). The lack of interdisciplinary research teams and funding has resulted in parallel literatures and disciplines that operate almost entirel independent of one another.

Fourth, although this has rapidly started to change, the survey and registration data mainly used by social scientists to study fertility has generally lacked any biomarker or genetic measures. Conversely, if combined genetic and register or survey data is available, many of the medical or genetic cohort datasets only include very crude measures of core social science and environmental indicators even though they are likely pivotal in understanding gene and socio-environment interactions (e.g., education, occupational or marital history, social capital and networks).

This reluctance, the lack of data, and the lack of interdisciplinary training that combines strong social and biological or genetic measures, has resulted in fertility theories and explanations in the social sciences that are generally socially deterministic, often solely based on concepts such as agency, motivation, conscious

choice and intentions (for review see Balbo et al., 2013), which are in turn highly conditioned by the environments of the family, peers, organizations, local and national institutional contexts (Mills & Blossfeld, 2005). Increasing evidence demonstrates that choice, agency and the behavioural outcomes that we examine in fertility, however, are not only socially-determined, but also linked to an individual's genetic architecture and beyond, such as proteins, hormones, neurons, gametes and other factors (Freese, 2008; Rietveld et al., 2014; Rietveld, Cesarini, et al., 2013; Rietveld, Medland, et al., 2013; Udry, 1996; Wachter & Bulatao, 2003; Wachter, 2008). Furthermore, recent research demonstrates that - as Udry (1996) and others hypothesized some time ago - a portion of the genetic influence of fertility is related to the motivational precursors to fertility (W B Miller et al., 2010; Rodgers, Kohler, et al., 2001). In other words, integrating genetics into the analysis of fertility will significantly improve our understanding and the explanatory power of our models. It is important to stress that biological explanations and heritability should not be mistaken for genetic determinism. As we describe shortly, the link between genes and phenotypic traits or outcomes may be indirect and mediated through environmental and psychological factors. This review will argue towards an integrative approach across disciplines.

1.1.3 The foundations: intermediate fertility variables/ proximate determinants

As a first attempt to integrate biological determinants within contemporary demography, Davis and Blake (Davis & Blake, 1956) introduced intermediate fertility variables, an approach which was later refined and quantified by Bongaarts (J Bongaarts, 1982; John Bongaarts, 1978) and Bongaarts and Potter (1983) under the label of proximate determinants of fertility. These seven factors that include both biological and behavioural determinants have been highly influential in shaping our theory and understanding of human fertility in demography – while remaining an exception of biodemographic work at that time. These are: (1) proportion of married women among all women of reproductive age (as a proxy to capture women exposed to sexual intercourse, also more broadly measured later by percentage of women in a sexual union, frequency of sexual intercourse), (2) contraceptive use

effectiveness, (3) duration of postpartum infecundability (or postpartum insusceptibility),[3] (4) induced abortion, (5) fecundability, (6) prevalence of permanent sterility; and, (7) spontaneous intrauterine mortality. The first four were considered to be the most important determinants due to the fact that they differ greatly between populations and the fact that fertility is the most sensitive to changes in these measures.

Criticisms and revisions of the model acknowledged the need to include sexual activities outside of marriage and a revised definition and estimate of total fecundity (Stover, 1998). Hobcraft and Little (Hobcraft & Little, 1984) also extended this work with their focus on fertility exposure analysis. These determinants were considered to be theoretically strong and highly plausible candidates for explaining the relationship between the level of fertility and both biological and behavioural factors. Yet further research using individual-level micro data demonstrates that additional factors also operate to explain the variation in fertility outcomes and levels (for review see: Balbo et al., 2013; Mills et al., 2011).

1.1.4 The emergence of a biodemographic approach to fertility

Recognition of the importance of biological or genetic determinants underpinning demographic behaviour in this area of research began to flourish in the late 1990s. Biodemography in general grew as a fruitful interdisciplinary approach, first applied in the area of longevity and mortality studies and the relationship between fertility-longevity interactions (Vaupel et al., 1998; Wachter, 1997, 2008).

Udry (1994; 1996) promulgated some of the earliest calls to include a biodemographic approach to study fertility and related behaviour in contemporary demographic research. His landmark article in 1996 went beyond the data constraints of that period and the application of simple behaviour-genetic models (Plomin, 1994) to hypothesize a series of probable biosocial relationships at the individual micro-level and societal macro-level. This included consideration of

[3] Bongaarts (1982) adapted the definition of postpartum infecundability to represent the combined effect of postpartum amenorrhea and abstinence.

additive, indirect and interaction effects between biological (hormonal, genetic) and social factors and environments. He likewise acknowledged that the behavioural choices or motivations to have children were likely guided by biological predispositions such as genetics, hormones, neurological structure or neurotransmitters. Such predispositions could be examined by proxy by personality for example, which has been shown as an antecedent to childbearing motivations and fertility outcomes (W B Miller et al., 2010; W B Miller & Pasta, 1994; W B Miller, 1992; Warren B Miller, 1994).

Another pivotal hypothesis introduced in Udry's (1996) paper outlined how changing social arrangements could alter the proportion of variance in fertility behaviour that is biologically controlled. In other words, in times when there are high normative and social constraints on the 'proper' timing of first birth and number of children a woman or couple should have, the less the variance in their social behaviour should be controlled by biological forces. Conversely, in the period of the second demographic transition where individuals and couples are less socially constrained and have considerable choice, biological forces should have a stronger influence on behaviour. Indeed, Udry's argument implies that biological predispositions should be more important now than in the past. This idea was adopted in further research to understand cohort differences, which we describe shortly (e. g. Briley, Harden, & Tucker-Drob, 2015; Kohler, Rodgers, et al., 2002; Kohler et al., 2006).

At the start of the 2000s, the biodemographic approach to fertility began to flourish with two interdisciplinary books appearing, *Offspring: Human Fertility Behaviour in Biodemographic Perspective*, co-edited by Wachter and Bulatao (2003) and *The Biodemography of Human Reproduction and Fertility* co-Edited by Rodgers and Kohler (Rodgers & Kohler, 2003). Both books provided broad coverage of the topic that goes beyond this limited review (e.g., endocrinology, neuroscience) and remain essential reading for this topic. Perhaps the most influential research and approach in fertility until now, however, has been the application of behavioural genetics models, which we turn to now.

1.1.5 Measuring genetic influence: family studies, twin studies and GREML models

The first questions to be answered in all behavioural genetics research are whether and to what extent genes have an influence on a trait[4] of interest (Guo, 2005). In other words, it is the basic question first asked by Kohler, Rodgers and Christensen in 1999: 'Is fertility behavior in our genes?' where they adopted a twin study approach to first ask and examine this question directly in demography, the findings of which will be described shortly.

The statistical concept used to measure genetic influences on a trait within a population is called *heritability*. Heritability is a population parameter (Visscher, Hill, & Wray, 2008) quantifying in how far differences between individuals are due to genetic differences. It is defined as the proportion of the genetic variance (V_G) over the overall variance of the trait (V_P phenotype) in a specific population (for details see Visscher 2008):

$$H^2 = \frac{V_G}{V_P}$$

Whereas the phenotypic variance is the sum of genetic and environmental variance.

$$V_P = V_G + V_E$$

The genetic component can be furthermore differentiated into additive (V_A) genetic effects—when the contribution of an allele to a phenotype is fixed—and non-additive (V_{NA}, epistatic and dominant) effects—when the contribution of an allele to the phenotype depends on the presence of other alleles or genes.

$$V_G = V_A + V_{NA}$$

H^2, defined as the genetic variance over the total variance in an outcome represents so called broad sense heritability. Generally, for most quantitative traits – a trait that shows continuous variation (Turkheimer, 2000) – it is assumed genetic variance is mainly due to additive effects . For fertility and other quantitative traits, empirical evidence for epistatic or dominant effects is negligible (Kohler et al., 2006;

4 This research often refers to a trait or phenotype, which in the social sciences is generally referred to as an outcome or dependent variable.

Zhu et al., 2015). Therefore, the genetic component underlying a trait is commonly quantified based on additive genetic effects and so called narrow sense heritability and when heritability is used in this volume it means:

$$h^2 = \frac{V_A}{V_p}$$

In other words, heritability quantifies the additive genetic variance component of an outcome variable.

It has been argued that all traits are heritable to some extent (Turkheimer, 2000) and that heritability is ubiquitous in social science (Freese, 2008). However, a number of caveats are in order. First, the genetic variance component provides insight into the extent to which genetic variation in a population is associated with variation in fertility, and a recent meta-analysis of all twin studies published up to 2012 has documented considerable variation in heritability across traits (Polderman et al., 2015). Second, fertility is a particularly interesting trait to study, because FTNS predicts that due to natural selection, there would be no heritability for traits such as NEB (Fisher, 1930). Nevertheless, we find a genetic component underlying fertility and it has been argued that the interplay of genes and the (social) environment is responsible for that (Hughes & Burleson, 2000; Kohler et al., 1999; Udry, 1996; chapters 2 & 4 are focussing on gene-environment interaction). Heritability can be estimated without measuring actual genes. For this reason, family (Bras, Bavel, & Mandemakers, 2013) and most commonly twin studies (Snieder, Wang, & MacGregor, 2010; see also chapters 2 & 5) are conducted. With the advent of molecular genetic data, a new analytical technique based on the genetic relatedness of unrelated individuals has become prominent (Visscher, Yang, & Goddard, 2010; Yang, Lee, Goddard, & Visscher, 2011; see also chapters 3 & 4; Yang et al., 2010), and we discuss these approaches in the next section.

Family Studies

Traditional family studies often follow simple designs and study e.g. parent-offspring correlations (Fisher, 1930; Pearson, Lee, & Bramley-Moore, 1899). Parents and children share on average 50 percent of their genes, so that two times the observed correlations in a trait, for example number of children ever born (NEB),

estimates the heritability assuming that the shared environment of family members does not play a role for. Correlations in fertility are very common among family members and also increased during the Twentieth century (Murphy, 1999). It should also be noted that additional fertility studies in demography have examined the impact of genes albeit indirectly via the study of the intergenerational transmission of the number of children (Anderton, Tsuya, Bean, & Mineau, 1987; Murphy & Wang, 2001), teenage pregnancy (Kahn & Anderson, 1992) and age at first birth (e. g., Barber, 2000; Steenhof & Liefbroer, 2008).

The pioneering work of Fisher (1930) found an intergenerational correlation of 0.20 in the NEB in data from around 2,000 British aristocrats born at the end of the Nineteenth century. He therefore estimated that 40 percent of the variance in the NEB within this population was attributed to genetic differences. The assumption that intergenerational correlations in fertility are entirely due to genetic effects, however, does not likely hold. In fact, we know from numerous studies in sociology and demography that shared environmental factors among family members or non-genetic (or cultural) inheritance of traits can lead to similarities in fertility among family members (e. g., Murphy, 1999; Steenhof & Liefbroer, 2008). The educational level or economic status, for example, are relatively stable across generations (Jæger, 2012; van Doorn, Pop, & Wolbers, 2011) and at the same time important for fertility outcomes (Mills et al., 2011). Therefore, it is necessary to separate genetic from environmental effects in families and within different societal contexts.

Twin Studies

The most common way to disentangle influences of latent family factors on any trait of interest is to use the information available in twin data. Twin models follow a relatively simple and straightforward logic and until now, represent one of the best resources for evaluating the importance of genetic variation in observed traits (Boomsma, Busjahn, & Peltonen, 2002).

Twin models facilitate the comparisons between two kinds of twins, identical or monozygotic (MZ) and fraternal or dizygotic (DZ) twins, in order to quantify genetic and non-genetic (or environmental) influences. In a classic twin design (as it is applied in chapters 2 and 5), MZ as well as DZ twins are siblings of virtually

identical ages and born and raised in the same families. The siblings are consequently assumed to share common environmental influences such as their parents, the neighbourhood they grew up in and other related aspects to the same extent, which is called the equal environment assumption. More importantly, MZ twins are genetically identical (i.e., share all their genotypes). DZ twins in contrast, are akin to full siblings and thus share on average only around 50 percent of their segregating genetic material. Similar to parent–offspring correlations, the correlation among DZ twins therefore reflects the importance of both environmental and genetic effects in families. The degree, however, to which MZ twins have a higher correlation in the trait of interest than DZ twins, reflects the degree to which they are genetically more similar. The comparison of twin correlations thus already makes it possible to quantify genetic and shared environmental effects on a particular trait (Boomsma et al., 2002; Snieder et al., 2010).

Since both kind of twins are assumed to share family environment to the same extent and MZ twins are twice as similar as DZ twins with regard to their genes, narrow sense heritability is two times the difference of the intra-group correlations of MZ and DZ.

$$h^2 = 2 * (\mathrm{r(MZ)} - \mathrm{r(DZ)})$$

The effect of the common environment of the twins (c^2) is therefore the pairwise correlation of MZ minus the heritability.

$$c^2 = \mathrm{r(MZ)} - h^2$$

Variance that is unexplained by these factors is due to non-shared environmental effects (e^2) from outside or even within the family; including measurement errors: for details (see Snieder et al., 2010).

Based on this logic, structural equation modelling (SEM) has become standard in twin research (as applied in chapters 2 & 5; see Boker et al., 2011 for details; Neale & Cardon, 1992). Above correlations have low power and large standard errors and do not make use of information available in variances and covariances. SEM furthermore provides goodness of fit statistics enabling the testing of alternative models and identifying assumptions – for example whether there are dominant genetic effects (Snieder et al., 2010). Extensions of the twin model to other relatives as well as other family designs are possible.

There has been some criticism of the potentially problematic assumptions and practical limitations of twin studies (see also chapter 2; for critical discussions on the equal environment assumption for example, see Horwitz et al. 2003). Both twin and family designs are likewise limited for other reasons in fertility research such as the fact that they require a pair of siblings, thus excluding singleton children and thus individuals from low fertility families. Since dizygotic twinning also has a genetic basis (Hoekstra et al., 2008; Mbarek et al., 2016), it is likely that these twins carry genes that might be important for high fertility. Thus, there might be a non-random genetic stratification which could bias variance estimates. Finally, these designs require multiple data on family-members, which are more difficult and costly to gather than data on unrelated individuals. With the advent of molecular genetic data, a new analysis technique has been developed to quantify genetic effects on a trait, namely genomic-relationship-matrix restricted maximum likelihood (GREML) models, which have been applied in chapters 3 & 4 and are discussed below.

GREML models

The GREML analyses (as applied in chapters 2 & 5) follow the same logic as the twin models. However, heritability estimates are based on the genetic similarity between pairwise unrelated individuals who do not share the same micro environment. To achieve this, a pairwise genetic similarity matrix is estimated from measured molecular genetic information across the entire genome. The approach capitalizes on recent developments in microarray technology to identify the association between specific traits and all common genetic variants across the whole genome. Such a design is possible due to the completion of the Human Genome Project in 2003 and the International HapMap Project in 2005 which provided the basic tools to find the genetic contributions of traits. As with other genetic data available until now, it is necessary to have the DNA from each participant in the study, often via a blood or saliva sample. Each person's DNA is then placed on tiny microarray chips and scanned on automated laboratory machines. These machines

quickly overview each person's genome for strategically selected markers of genetic variation, referred to as SNPs (single nucleotide polymorphisms[5]).

The GREML analyses correlate the genetic similarity with the phenotypic similarity between individuals. More precisely, based on a mixed linear model, it performs a restricted-maximum likelihood (REML) analysis to decompose the observed variance of an outcome into genetic and residual variance (environmentally caused and measurement error). The method has first been applied to height as a model complex trait (Visscher et al., 2010; Yang et al., 2010). Chapters 3 and 4 exploit these (quite) recent advances in the field of molecular and quantitative genetics. Note that in contrast to genome-wide association studies (GWAS, see 1.4.1) GREML methods do not aim to identify individual genetic variants associated with an outcomes, but to quantify the genetic variance explanation of an outcomes based on all common SNPs from the whole genome.

1.2 The chapters of this volume in the light of previous genetically informed fertility research

This section summarises a series of studies that have focussed on estimating the heritability of NEB and AFB, using the aforementioned behavioural genetics approach. Gene and environmental interaction at the macro-level is then touched upon, concluding with multivariate models that go beyond heritability estimates and a brief mention of topics related to motivation and education. We localize and detail the contributions of the chapters in this volume in the light of the literature (for an overview of the chapters in this volume see Table 1-1, section 1.3).

1.2.1 Heritability studies

Several twin and family studies provide evidence for a genetic component underlying both the tempo (AFB) and quantum (NEB) of human fertility up to a

5 DNA consists of around 3 billion pairs of nucleotide molecules and each can be indexed by its location on the genome. There is no variation for most of these nucleotides across individuals; rather, we differ in our polymorphisms and hence the focus on SNPs, which occur once every 100 to 300 nucleotides. A single gene can contain hundreds of SNPs.

level of 50 percent (e. g. Byars, Ewbank, Govindaraju, & Stearns, 2010; Kirk et al., 2001; Kohler et al., 1999; Milot et al., 2011; Neiss, Rowe, & Rodgers, 2002; Nisén, Martikainen, Kaprio, & Silventoinen, 2013; Tropf, Barban, et al., 2015; Zietsch et al., 2014). Figure 1-1 provides an overview of selected studies, which investigated the heritability of fertility behaviour across birth cohort. The figure shows the estimates for women for age at first birth (AFB) and for both men and women of number of children ever born (NEB). Here we can see that heritability estimates for AFB for women range between 0.001 (USA 1920-1935 and 1956-1970) and 0.35 (UK 1930-1939) or in other words that up to just over 0 or 35 percent of the observed variance in AFB within these birth cohorts is due to additive genetic effects. NEB in women ranges from 0.01 (USA 1920-1935) to as high as 0.59 (USA 1936-1955) and for men, the range for the same trait is between 0 (Denmark 1880-1900) and 0.56 (Netherlands 1810-1820). The figure shows that there is considerable variation in heritability estimates across time, between countries and the sexes, a point we return to later.

Heritability of AFB

The age at first birth (AFB) has been assessed in a recent study by Nisén, Martikainen, Kaprio, & Silventoinen (2013) on Finnish twins born between 1950 and 1957. For men they found that common environmental factors between siblings – and not their genes – mainly explain resemblance in the AFB among brothers. For women, in contrast, they estimate heritability of the AFB at 0.26, which is also in line with studies from Australia (Kirk et al., 2001) and the U.S. (Byars et al., 2010 as well as our findings in Chapter 2). However, the AFB turns out to be a challenge to study, since other studies, e.g. in the U.S. (Neiss, Rowe, & Rodgers, 2002) and Denmark (Rodgers et al., 2008), show no significant effect on the AFB of women.

Chapter 2 in this dissertation investigates for the first time genetic effects on the AFB among twins from the UK born 1919-68. A core issue in the analysis of AFB is the handling of right-censored observations (i.e., individuals that did not have a first birth by the last time of observation). In contrast to the commonplace practice of including right-censored cases in demography and sociology (e.g., Mills 2011), it is uncommon to deal with censored information in genetic studies and childless women are generally excluded from the sample (Byars et al., 2010; Neiss et al., 2002;

Nisén et al., 2013; Rodgers et al., 2008). Using data from the TwinsUK cohort (Moayyeri, Hammond, & Spector, 2013), Chapter 2 empirically examined whether the inclusion of censored information in a Tobit model affects heritability estimates compared to the classic models. Results show strong and non-systematic differences between both kinds of models suggesting that the comparison of these research designs has to be critically reconsidered. The Tobit model estimates – in line with previous studies from Australia and Finland – that a substantial part (40 percent) of the variation in AFB is associated with latent family characteristics. Genetic dispositions (26 percent) play a more important role than the shared environment of siblings (14 percent), with the nonshared environment/measurement error having the strongest influence (60 percent).

In Chapter 3, we extended this analysis and applied the recently developed GREML method, which is based on molecular genetic data and uses the genetic similarity of unrelated individuals, who share no micro environment to estimate the heritability of AFB. Given that this method makes use of only the fraction of genetic variance amongst unrelated individuals, large sample sized are required and we combined data from the TwinsUK with the Lifelines Cohort Study (Klijs, Scholtens, & Mandemakers, 2015) to increase statistical power. The analysis based on around 5,000 individuals estimates SNP-based heritability to be 0.15, meaning that common SNPs explain 15 percent of the observed variance in the AFB across individuals in the combined sample.

Heritability of NEB

Focusing on NEB, a study by (Rodgers, Kohler, et al., 2001) investigated genetic and environmental influence on the NEB among female and male twins from Denmark born between 1953 and 1959. They found that twin correlations among MZ pairs (~0.30) are more than twice as high as those of DZ twin pairs (0.13) for both males and females. The structural equation models estimate heritability for both sexes at around 0.30, meaning that 30 percent of the observed variance in the NEB within these birth cohorts is due to additive genetic effects. Both non-additive genetic and shared environmental effects were not significant, and therefore the remaining 70 percent of the variance can be attributed to non-shared environmental

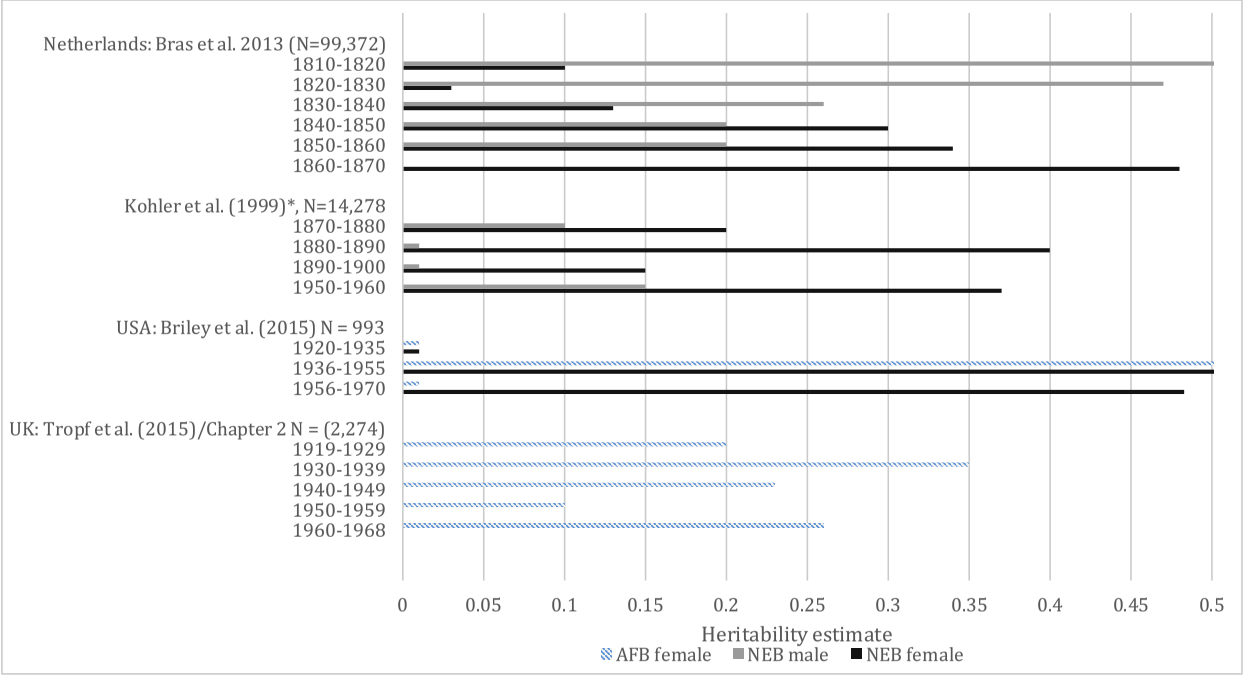

Figure 1- 1 Summary of fertility heritability estimates of selected studies by birth cohort and country.

Note: AFB = age at first birth, NEB = number of children every born, *= in the study by Kohler et al 1999, we looked up the approximate estimate from Figures 2,3 – for males based on a dominant genetic model.

effects between the siblings, which also includes measurement error. These findings are almost perfectly replicated in a recent study of Swedish twins born between 1915-1929 (Zietsch et al., 2014). Not only fertility research, but behavioural genetic research in general shows that considerable variation is attributable to non-shared environmental factors. Unique environmental factors appear to be the main source of variation; amongst others, the partner of the individuals plays an important role for fertility behaviour (Kohler & Rodgers, 2003).

In chapter 3 of this dissertation, we also investigated genetic influence on NEB in the the GREML model. We found that 10 percent of the variance in the NEB is explained by genetic factors – as for AFB, this is smaller than typical estimates from twin models. In general, it is not unusual that heritability estimates obtained by GREML utilizing molecular genetic data are lower compared to estimates from family or twin studies. Yang, Manolio et al. (2011) observed that such common SNP heritabilities cover one third to one half of the total heritability found in twin and family studies with the remaining genetic variance most likely explained by lower frequency variants (Visscher, Brown, McCarthy, & Yang, 2012; Yang et al., 2015). Another possibility is that non-additive genetic factors (Zhu et al., 2015; Zuk & Hechter, 2012) and/or gene-environment interaction inflate heritability estimates from twin studies (see chapter 4).

1.2.2 Beyond heritability estimates: GxE interaction and multivariate models of fertility behaviour

Gene-environment interaction (GxE) refers to the situation where genetic effects that are associated with a trait are dependent upon the variability in the environment or when genes modify an organism's sensitivity to particular environmental features (e. g. Shanahan & Hofer, 2005). A particular context can trigger or compensate for a genetic vulnerability and improve adaptation. In general differences in heritability estimates between birth cohorts and countries apparent in Figure 1-1 may suggest gene-environment interaction in reference to fertility. These findings suggest that biodemographic research on fertility needs to continue to embrace demographic and sociological research, because different heritability

estimates for a trait between different populations do not necessarily point purely to methodological issues, but can be informative in substantive terms.

As described previously, researchers such as Udry (1996) hypothesized that there is likely an interaction of biological with societal level factors at the macro- or population level. In fact, one of the most common misconceptions about genetic studies is that the heritability of a trait within one population can predict the heritability in the same trait in another population – even though similarities in estimates can be remarkable for physical traits such as height (Visscher et al., 2008). First, heritability is a ratio of genetically caused variance over the overall variance. Therefore, changes in the overall variance in a trait can change heritability independent of the genetic variance component, or in other words – heritability is a population specific parameter. Second, the genetic variance component depends on the genetic endowment, for example, of allele frequencies within a population. Third, and presumably most important for fertility, genes always only encode predispositions for a trait, but environmental conditions determine whether these dispositions become manifest (Guo, 2005). Thus, genes and environment can interact and replication in different temporal and spatial settings is pivotal to gain insight about this interplay.

Gene and environment (GxE) interaction in fertility across populations and birth cohorts

Gene x environment (GxE) interaction has been examined in the study of NEB. Using data on the historical Danish twin cohorts including virtually every twin pair born since 1870, Kohler et al. (1999) and Kohler, Rodgers, & Christensen (2002) report large changes in heritability estimates of the NEB for cohorts born during the past centuries. Particularly, after the strong fertility decline of the First Demographic Transition at the end of the nineteenth-century and the Second Demographic Transition in the second half of the twentieth-century, heritability reached a moderate level of 0.40 – while getting close to zero in the interim phases.

As described in more detail previously, this fits with the explanation hypothesized by Udry (1996) who described changes in the influence of genes on reproductive behaviour in terms of GxE with factors such as societal norms

constraining choice and behaviour. Udry argued that genetic predispositions for fertility gain importance in environments that are less restrictive in their fertility choice. At the same time, genetic predispositions play a minor role in restrictive social contexts such as strong normative rules about the timing and number of children, war or economic crisis.

Following Udry's (1996) hypotheses, Kohler et al. (1999) associated the observed peaks in heritability with the environmental changes during the Demographic Transitions. In particular, there were improvements in economic, medical and hygienic conditions during the First Demographic Transition, with the Second Demographic Transition characterized by the introduction of the pill and cultural transformations relaxing fertility norms triggering genetic expressions (van de Kaa, 1987). A recent study corroborates this reasoning, applying an extension of the family model on a large scale database containing 100,000 sibling pairs from the Dutch Province of Zeeland born between 1810-1870 (Bras, Van Bavel, & Mandemakers, 2013). These findings not only show for women a rise in heritability during the first Demographic Transition, but they present evidence that this rise was especially pronounced for women in more urban areas or with a liberal religious climate, thus underscoring the idea that individual freedom triggers genetic influences on the NEB.

In Chapter 2 of this dissertation, we analysed changes in the genetic influence on the AFB across birth cohorts from 1919-1968 of female UK twins. We also found large swings in heritability co-varying with strong environmental upheavals. While heritability is low or non-significant for birth cohorts who started childbearing during the end of World War II or the economic crisis of the 1970s and 1980s, it rises to around 0.40 for individuals who started childbearing in times of changing mores and the sexual freedom in the 1950s and 1960s. These results appear to be well in line with previous studies on both NEB (e. g. Bras et al., 2013; Kohler et al., 1999) and AFB (Briley et al., 2015), confirming the initial rationale of Udry (1996).

It is important to note that Udry's heuristic assists interpretation of the observed trends in heritability estimates over time and across groups. However, estimates from twin studies typically have large standard errors, and while differences across time (Kohler et al., 1999; Tropf, Barban, et al., 2015) or religious groups (Bras et al., 2013) are in the expected direction, they remain mostly non-

significant. Additionally, previous studies have reported findings of substantial heritability (0.30-0.50) in traditional, pre-transitional societies (Milot et al., 2011), for which the FTNS and Udry's theory would predict low heritability. Ongoing and future research based on molecular genetic data, particularly polygenic scores (see Conley & Domingue, 2016; Conley et al., 2016 for applications), will permit testing of specific hypothesis in statistically more powerful ways that twin- and family studies.

In chapter 4, we went one step further than than twin studies and investigated whether the same genes are important for NEB and AFB in different geographical and temporal environments. We conducted a mega-analysis of individual level data—in contrast to a meta-analysis of summary statistics—applying whole genome methods on 31,396 unrelated men and women from six Western countries (Australia, Estonia, Netherlands, Sweden, UK, US). Across all individuals and environments, common SNPs explained only ~4 percent of the variance in NEB and AFB. We then extended these models to test whether genetic effects are shared across different environments or are 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 percent for NEB and 19 percent 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.

Multivariate modelling

Although the focus until now in this review has been on NEB and AFB, other related fertility traits have also been examined. Genetic influences have been explored for the age at first dating or marriage (Mealey & Segal, 1993), age at first sexual intercourse (Dunne et al., 1997), number of sexual partners (Guo, Tong, & Cai, 2008) and likelihood of unprotected sexual intercourse (Daw & Guo, 2011). One of the strongest genetic effects was found in a twin study examining the first attempt to

have a child among the Danish twins, measured in retrospective interviews of the first attempt to conceive. In this study, up to 50 percent of the variance in that trait was explained by genetic differences (Kohler et al., 1999; Rodgers et al., 2001). This raises the question of how different fertility traits relate to each other and to what extent the same genes or the same environments are important for different fertility traits.

Several studies engage in more complex (twin) modelling approaches (e.g., Kohler, Behrman, & Schnittker, 2011) estimating, for example, bivariate genetic models. Following the same logic as in classic twin studies, it is possible to estimate the extent to which the same genetic and environmental factors are important for two different traits. If trait 1 of twin 1 predicts trait 2 of twin 2, the correlation of traits runs in families and therefore there is a common etiology of the traits within the family. If this cross-trait-cross-twin correlation is higher for MZ twins than for DZ twins, the common etiology is partly genetic. If the genetic correlation between two traits is 1, all genetic variance in trait 1 and 2 has a common base. If the genetic correlation is 0 the genetically based variance between trait 1 and 2 are independent. For a more detailed empirical explanation, refer to Kohler et al., (2011).

The age at first birth and number of children ever born

Two studies from Australia (Kirk et al., 2001) and the US (Byars et al., 2010) have used twin and family designs to investigate whether genetic influences on the AFB and NEB have a common base. Both studies find that the well-established negative correlation between AFB and NEB is partly genetic, meaning that genes that are associated with a later AFB are also associated with a lower NEB across the life-course. In chapter 3, we also investigated whether genetic effects on AFB and NEB are correlated, but did this using bivariate GREML models, which produce unbiased estimates of shared genetic effects amongst unrelated individuals (S. H. Lee, Yang, Goddard, Visscher, & Wray, 2012; Visscher et al., 2014). In the aforementioned sample of women from the Netherlands and UK, we find a statistically significant genetic correlation between AFB and NEB of –0.62 (SE = 0.27, p-value = 0.02), implying that a considerable part of the underlying genetic factors are shared between AFB and NEB.

Decision-making and fertility motivations

One area of research has focused on decision-making and motivations. Rodgers et al. (2001) combined information on the NEB in the Danish twin cohorts with the first attempt to become pregnant. They found firstly, that the age at first attempt to conceive is heritable for both men (0.35) and women (0.52) and secondly, that around 10 percent of the genetic variance in the NEB is related to the genetic variance in the first attempt to conceive. This study can be seen as key evidence that part of the genetic dispositions influencing our realized fertility is not due to physical traits, but mediated by conscious decision making.

A more recent study by Miller et al. (2010) empirically demonstrated that a portion of the genetically influenced fertility behaviour is related to the motivational precursors to fertility. Examining both timing and NEB and using the NLSY79 in the U.S., they conducted both uni- and multivariate behavioural genetic analyses. The central finding was that there are additive genetic effects that operate through desires to have a certain number of children that affect both the timing of the next child that is born and the ultimate NEB. They also link these affects to gender role attitudes but find no genetic variance associated with either child-timing intentions or educational intentions. This was an extension of previous work such as that by Miller (1994) and Pasta and Miller (2000), showing that both positive and negative childbearing motivations were significantly heritable and that the fertility outcomes we observe are results of prefernces. This builds on Miller's (1994) Traits-Desires-Intentions-Behaviour (TDIB) theoretical framework.

Education, fertility and causality

Education – and particularly women's education – is a central predictor of particularly the timing of first and subsequent births (Mills et al., 2011). Multivariate genetic models have been applied to explore the relationship between education and fertility with respect to its common genetic and environmental base.

The aforementioned study on Finnish twins (Nisén et al., 2013) shows that the negative association of education with the chance of having any children among women is largely genetically based. The same reasoning accounts for the positive association between education and the chance to have any children for men, suggesting that different genes are important for men and women. Using more complex models, two studies from the U.S. (Neiss et al., 2002) and Denmark (Rodgers et al., 2008) test whether education mediates the positive relationship between intelligence and AFB of women in a behavioural genetic framework. Results show that education partially mediates the effect of intelligence on AFB in standard models, but that this mediation is not significant once a model is fitted that takes latent family factors (environmental and genetic ones) into account. In other words, differences in intelligence, education and the AFB emerge from variance between families, not differences between siblings within the same family.

Such findings may cast doubt on the assumption that education causally affects the timing of childbearing and thus has the potential to challenge the widely accepted assumption that educational expansion of women is the main explanation for the massive postponement in the AFB during the second half of the Twentieth Century (Balbo et al., 2013; Mills et al., 2011; Ní Bhrolcháin & Beaujouan, 2012). In chapter 5, we focus on this question and apply within-twin regression models as well as multivariate behaviour genetics models on the association between education and age at first birth in a large sample of twins from the UK. Within-twin models include fixed effects per family, which capture all (observed and unobserved) factors shared among the twins – including genetic effects. The estimated effect of education on the age at first birth can therefore be interpreted as a causal effect (for a critical discussion see Amin, Behrman, Kohler, Xiong, & Zhang, 2015; Boardman & Fletcher, 2015). Kohler and colleagues (2011) combined this approach with the behaviour genetics structural equation models (the so called ACE-beta model), to investigate simultaneously the causal effect of education on fertility and the effect of genes and the shared (social) environment of siblings on the education-fertility link.

Our findings show that one year of additional schooling is associated with about half a year later age at first birth in standard models. This fell to only 1.5 months for the within-identical twin model that controls for all shared family

background factors (genetic and family environmental). The ACE-beta model revealed that mainly influences of the family environment – not genetic factors – cause spurious associations between education and age at first birth. Lastly, we demonstrated using data from the Office for National Statistics that only 1.9 months of the 2.4 years of fertility postponement for birth cohorts 1944-1969 could be attributed to educational expansion based on these estimates. We conclude that (the rise in) educational attainment alone cannot explain differences in fertility timing (between cohorts).

1.2.3 Evolutionary Anthropology and Biology

The behavioural genetics modelling of human fertility has obviously not been an exclusive interest of social scientists, but has also been widely applied by evolutionary biologists and anthropologists. The question of evolution was initially more of a question of surviving until one reproduces, since only those who survived were able to transmit their genes to the next generation. However, due to improvements in hygiene and the subsequent reduction in prenatal, infant and child mortality, whether and how many children an individual bears has become are important question in understanding evolution. If genetic variants are associated with fertility success, this means that some genetic variants will be more frequent than others in subsequent generations, therefore indicating natural selection and evolution (Stearns, Byars, Govindaraju, & Ewbank, 2010).

This line of reasoning, however, theoretically predicts that fertility as well as other behaviour important for the transmission of genes to the next generation show low or no significant genetic variation (Kimura, 1958) within a population, because evolution should have already trimmed the differences over time. As mentioned at the start of this review, it has been assumed that this so-called Fundamental Theorem of Natural Selection (FTNS, Fisher, 1930) is the reason why genetically informed fertility research has been neglected for many decades (Kohler et al., 1999; Rodgers et al., 2001).

However, empirically this FTNS turns out to be only partly true. On the one hand, we do observe a much higher heritability for morphometric traits (for example height: ~ 0.80 in humans) than for traits such as fertility (Visscher et al., 2008). On the

other hand, fertility traits do still show significant heritability. Kirk and colleagues (Kirk et al., 2001), for example, analysed the heritability of the NEB to determine to what extent genes influence the relative reproductive advantage in a population of Australian female twins. They find a heritability of 0.36, indicating that modern societies are still under natural selection.

Evidence for natural selection

The genetic correlation of a trait with NEB furthermore indicates whether a specific trait is under natural selection. Findings from the Framingham Heart Study, for example, suggest that the same genes influencing the NEB are negatively correlated with a number of traits, among others height, cholesterol levels, systolic blood pressure and the AFB (Byars et al., 2010). Consequently, the authors expect selective changes in the disposition for these traits for subsequent generations (e.g., that the female US population under study will shrink in body-size).

The negative association between NEB and the AFB leads to a prediction of a decrease in the AFB across generations (Byars et al., 2010; Kirk et al., 2001). Such findings are in line with studies on natural populations such as from Milot and colleagues (2011). They observed in a historical population of natural fertility (i.e., the population of ile aux Coudres with high fertility norms in a stable environment) in Canada that the AFB decreased by around 4 years within the past 140 years as a response to natural selection. Kirk and colleagues (2001) argue that in contemporary times, as compared to the past, more psychological and social traits are important for NEB and AFB, and are therefore assumed to be under natural selection. A recent investigation by Stulp and colleagues (2015) shows that in the Netherlands, height in men was consistently positively associated with NEB throughout the past century, which may partly explain the fact that the Dutch are the tallest people on earth.

Our findings of additive genetic variance in NEB (chapter 3) thus also suggest the potential for on-going natural selection on AFB. The genetic correlation between AFB and NEB furthermore suggests that women with a predisposition for an earlier AFB will have had more children and therefore more frequently passed on their genes to the next generation. From an evolutionary perspective, we would expect a shift in the AFB to younger ages in more recent birth cohorts. However, what we observe is the exact opposite, namely, a massive postponement in the AFB in most

recent birth cohorts across developed societies (Balbo et al., 2013; Mills et al., 2011). The same has been reported for educational attainment. While the negative genetic correlation between educational attainment and fertility in the US would predict a decrease in educational attainment on the population level across generations, the average years spent in education have steadily increase throughout the past century (Beauchamp, 2016). It is therefore important to acknowledge that environmental influences such as the introduction of the 'pill' or political reforms are much more relevant for trends in behavioural outcomes such as fertility than natural selection might be (for a detailed discussion see Courtiol, Tropf, & Mills, 2016).

1.3 Summary table of chapters in this volume

Table 1-1 Summary Table of Chapters in this Volume

Chapter 2: Genetic influence on age at first birth of female twins born in the UK, 1919-68			
Questions Do genes influence the age at first birth in the United Kingdom? Have genetic influences on AFB changed during the past century? Are there differences in the estimates when in- or excluding childless individuals?	**Outcomes** The age at first birth; the age at first birth including childless individuals with their last year of observation. **Predictors** Additive genetic effects (A), shared environmental influences of twins (C), unique environmental influences on an individual (E).	**Data** Female twins from the United Kingdom born between 1919-68 from the TwinsUK register; N = 4,548. **Methods** ACE Structural equation modelling (SEM); ACE tobit models in the SEM framework (including childless individuals)	**Results** Genetic dispositions explain 26%, the shared environment of siblings 14% and the nonshared environment/measurement error 60% of the differences in AFB. Estimates change over time, we find support for the idea that environmental constraints (war or economic crisis) suppress, and normative freedom (sexual revolution) triggers the activation of genetic predispositions that affect fertility. The exclusion of censored information (i.e. on the childless) changes results non-systematic.

Chapter 3: Molecular genetics, human fertility and natural selection in modern societies			
Questions To which extent do common genetic variants influence the a) age at first birth b) number of children ever born of women from the UK and Netherlands? Do genetic effects on the age at first birth and the number of children overlap?	**Outcomes** Age at first birth; number of children ever born. **Predictors** Common SNPs/ genetic relatedness matrix estimated from 1,017,420 SNPs. **Controls** First 20 principal components of genetic relatedness matrix, year of birth, zygosity, country	**Data** One in a pair of female twins from the United Kingdom from the TwinsUK register. Dutch women from the Lifelines Cohorts Study (genotyped subsample); N = 6,758. **Methods** Univariate and bivariate genomic-relationship-matrix based restricted maximum likelihood methods (GREML).	**Results** Significant additive genetic effects on both traits explain 10% (SE = 5) of the variance in the NEB and 15% (SE = 4) in the AFB. Genetic effects on both outcomes show a negative genetic correlation between AFB and NEB of –0.62 (SE = 0.27, p-value = 0.02). This finding implies that individuals with genetic predispositions for an earlier AFB had a reproductive advantage and that natural selection operated not only in historical, but also in contemporary populations. The contrasting observed postponement in the AFB across the past century in Europe suggests environmental effects to be stronger.

Chapter 4: Mega-analysis of 31,396 individuals from 6 countries uncovers strong GxE interaction for human fertility			
Questions Are the same genes associated with number of children ever born or age at first birth a) across geographical b) and historical environments?	**Outcomes** Age at first birth; number of children ever born **Predictors** Common SNPs/ genetic relatedness matrix estimated from 847,278 SNPs. **Controls** First 20 principal components of genetic relatedness matrix, year of birth, zygosity, country	**Data** We combined data from the Atherosclerosis Risk in Communities Study (ARIC), the Health and Retirement Study (HRS), the Estonian Genome Center, University of Tartu (EGCUT), the Queensland Institute for Medical Research in Australia (QIMR), Lifelines Cohort Study in Groningen, the TwinsUK and the Swedish Twin Register (STR); N= 31,396. **Methods** Univariate and bivariate multiple genomic-relationship-matrix based restricted maximum likelihood methods (GREML).	**Results** Across all individuals and environments, common SNPs explained only ~4% of the variance in NEB and AFB. 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. 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.

Chapter 5: Is the association between Education and fertility postponement causal? The role of family background factors			
Questions Is there a causal effect of education on the age at first birth? To what extent do genes and the family environment and genes cause spurious associations between education and the age at first birth? How much of the observed incline in the age at first birth during the second half of the Twentieth century can be explained by the educational expansion?	**Outcomes** Age at first birth; the age at first birth including childless individuals with their last year of observation **Predictors** Age when leaving full-time education, additive genetic effects (A), shared environmental influences of twins (C), unique environmental influences on an individual (E), year born.	**Data** Female twins from the United Kingdom born between 1919-69 from the TwinsUK register; N=2,752. Office for National Statistical in the UK: General Household Survey 1935-70 (N=35,435) for education. Aggregate statistics on the age at first birth. **Methods** OLS regression models; within twin pair fixed-effects regression models; bivariate ACE Structural equation modeling (SEM); bivariate ACE-beta Structural equation modeling (SEM)	**Results** While standard models showed that leaving full-time education one year later is associated with about half a year later age at first birth this reduced to only 1.5 months for the within-identical twin/ causal model that controls for all shared family background factors (genetic and family environmental). Biometric analyses reveal that mainly influences of the family environment – not genetic factors – cause spurious associations between education and age at first birth. We demonstrate that only 1.9 months of the 2.4 years of fertility postponement for birth cohorts 1944-1969 could be attributed to educational expansion based on these estimates.

1.4 Conclusions and discussion: Limitations and fertile future research frontiers

The aim of this introductory chapter was to examine the current state of knowledge and place the contributions of this book into context withing the biodemography literature. After touching upon the different terminology and underlying reasons for the lack of research into this area, we turned to the foundations of approaches that included both behavioural and biological determinants of fertility. The current review summarized findings from behavioural genetic research on this topic which show in twin and GREML models that there is clearly a genetic component to fertility outcomes, with both AFB and NEB having a heritability ranging from zero to levels of just over 40 percent (as examined in chapters 2 & 3). Since heritability is a population parameter, heritability of a trait within one population cannot predict the heritability within another (see Visscher et al. 2008, see chapter 4 for a demonstration). This leads us to acknowledge the importance of the environment and the promise of studies that have examined swings in heritability with environmental changes and choice in fertility decisions. Other fruitful directions are studies that have gone beyond heritability estimates to probe into more complex multivariate models of fertility behaviour such as first attempts to conceive a child (Rodgers et al. 2001), motivational precursors to fertility (Miller 1994; Pasta & Miller 2000; Miller et al. 2010) and the (causal) relation between education and the AFB (see chapter 5).

1.4.1 Genome-wide association studies (GWAS)

One central contribution of this dissertation was to introduce molecular genetic data to the study of age at first birth and number of children ever born (chapters 3 & 4). Virtually all of the fertility-related research conducted within biodemography had adopted a behavioural genetics twin approach until now. The main advantages of using molecular genetic information and GREML approaches

are: first, that heritability estimates are unbiased by potential gene-environment correlations in families; and second (chapter 4), that we were able to relate the genetic material of individuals across countries and birth cohorts to investigate whether genetic effects on age at first birth and number of children ever born are the same across environments.

Recent advances in molecular genetics, however, offer new possibilities. Simply put, the main distinction is as follows: whereas the models we presented focused on whether fertility is in the genes and to what extent it is heritable, molecular genetics may attempt to isolate where the specific genetic variants responsible are located and subsequently focus on the function of those genes identified.

Advances using a molecular genetic approach first adopted a candidate-gene approach, which has been criticized and requires further replication, but provides an initial window into possible mechanisms and approaches that might be interesting to pursue in future research (Daw & Guo, 2011; Guo, Tong, & Cai, 2008; for critique see Hewitt, 2012). More promising approaches are hypothesis-free methods to find genetic variants related to AFB and NEB via a Genome Wide Association Study (GWAS). Since it is highly implausible that even the best social scientist could specify the biological pathway and isolate particular genetic variants in advance, this approach offers an alternative.

GWASs emerged as a promising new approach to connect genetic variants to the outcomes of interest (Zhao & Chen 2013). GWAS refers to hypothesis-free testing without any a priori assumptions about either the biological pathway or a particular location (Nolte et al. 2010). It likewise embraces the fact that there are multiple genes (polygenic) and pathways associated with fertility that are difficult to specify in advance with our current state of knowledge. A GWAS therefore runs millions of separate regressions on the phenotype (outcome) of interest across the entire genome. Due to the large number of SNPs that are tested in GWAS, an association must achieve a stringent threshold of statistical significance ($P < 5x10^{-8}$) in order to be considered valid. An association identifies a genomic region and not a specific causative mutation that might be involved in the development of the trait or behaviour.

The computational GWAS approach remains promising for social science genetics due to the fact that it overcomes some of the mistakes inherent in candidate-gene studies in the past and due to the fact that the often elusive biological pathways underlying the trait do not need to be specified in advance. It is also the only technique currently available that has the potential to discover novel genes which could then be used in further more reliable follow-up studies and provide directions to researchers where they need to search and which potential biological pathways to pursue. It also allows population stratification to be controlled, which remains a key issue in avoiding bias and misinterpretation of results in this type of research (see Wray et al. 2013 for a review).

A promising first application in the social sciences examined genetic variants that predict educational attainment on more than 120,000 genotyped individuals by Rietveld, Medland et al. (2013) as part of the Social Science Genetics Association Consortium (www.ssgac.com). At the time of writing this review, Mills and her research team at the University of Oxford are currently leading a large consortium to engage in the first ever GWAS and meta-analysis of reproductive choice (age at first birth; number of children), conducted in both men and women in over 60 data sets.

Although GWAS are able to narrow down where to look in the genome, recent research has retorted (Hirschhorn 2009) that the validated SNP associations explain only a small portion of the genetic variance or heritability of the phenotype that has been estimated from classic family and twin studies; this is often referred to as the 'missing heritability' problem (Manolio et al., 2009). Despite the fact we are currently working towards isolating SNPs robustly associated with reproductive behaviour in the previously mentioned GWAS, we show in chapter 4, based on the same type of molecular information utilized in GWAS that this is not the entire story of the genetic architecture of reproductive outcomes. Instead, for a complex outcome such as reproductive choice, we will be able to explain some – but not all – of the genetic heritability with a GWAS. Instead, it appears that the interaction of genes with the environment are likewise crucial.

Although no GWAS has been published as of yet, among the reproductive choice variables we often study in demography and sociology, several studies have shown promising results for fertility-related outcomes related to infertility and the

reproductive life span (age at menarche, age at menopause). Previous research has successfully demonstrated that there is a genetic component to reproduction with over 70 genome-wide association studies (GWAS) published for 32 traits and diseases associated with reproduction (see Montgomery et al. 2014). This includes identification of genes such as those related to age at menarche (Elks et al., 2010; Sulem et al., 2009), age at menopause (Day, Ruth, & Thompson, 2015; He et al., 2009; Perry et al., 2009; Snieder, MacGregor, & Spector, 1998; Stolk et al., 2009, 2012), and endometriosis (Painter et al., 2011).

The results from a recent study among a natural fertility population of Hutterites by Kosova et al. (2010) examining male fertility traits and isolating 41 SNPs, likewise shows promising areas to examine further. In this study, nine of the SNPs were associated with reduced sperm quality, providing a further link to potential biological pathways in men. Due to a strict religious doctrine in the Hutterite community, the variation in non-genetic factors is minimized between individuals, allowing them to confirm the presence of a significant genetic component in the natural variation of fertility. The use of these types of results, particularly in interaction with the socio-environment, could lead to a new understanding of fertility as we know it within the social sciences.

1.4.2 Polygenic scores (PS)

The genetic architecture of many human traits appears to be influence by many genes and is therefore polygenic (Visscher et al., 2012). Previous experiences in social science genetics have lead to the conclusion that a small number of SNP cannot explain a large proportion of variance in an outcome, but that: "A typical human behavioural trait is associated with very many genetic variants, each of which accounts for a very small percentage of the behavioral variability" (Chabris, Lee, Cesarini, Benjamin, & Laibson, 2015, p. 305). As mentioned before, existing evidence on the genetic architecture of reproductive traits such as age at menarche and age at menopause also suggest an underlying polygenic structure. Scholars have proposed creating an additive polygenic score (PS) depending an p-value thresholds from GWAS findings and potentially weighted by effect sizes (Purcell et al., 2009).

Such scores had considerable success in predicting complex and quantitative outcomes such as schizophrenia (Purcell et al., 2009) or body mass index (BMI) (Speliotes, Willer, & Berndt, 2010).

Utilizing findings from GWAS studies, such techniques have at least three great advantages for genetically informed social science fertility research. First, if the predictive power of a PS for fertility is sufficiently large, the PS could be used to control for genetic endowment in social science studies, which otherwise might represent unobserved heterogeneity and bias the estimates of other social science predictors in a statistical model (Rietveld, Medland, et al., 2013). Second, with the integration of PS in social science research it becomes possible to approach the biological and social pathways of genes on fertility behaviour by, for example, predicting fertility from education PSs or other related social science outcomes (Beauchamp, 2016; Conley et al., 2016) or test for mediation of fertility PSs in multivariate regression models. Third, it may give new insights into causal pathways amongst social science outcomes. A recent study on the relationship between parental age and schizophrenia, Mehta et al. (2016) show that the well-established association between parental age and children's well-being might be partly genetic in nature. So far, the association had been attributed to unstable relationships of young parents and their low economic resources as well as the decreasing (physical) health of older parents which, for example, may complicate conception, pregnancy and birth (Goisis, 2013). Mehta et al. find a U-shaped relationship between the PS for schizophrenia and age at first birth, that is, genes for schizophrenia are associated with relatively old (>40) or young (<20) age at first birth. Given that genes influence parents age at first birth as well as their children's mental health via inheritance, the intergenerational association might partly be spurious and not considering genetic effects may bias findings in social science studies.

1.4.3 Towards truly interdisciplinary work and quality control

A central challenge to adopting a biodemographic approach in fertility research has been the lack of training and understanding of very different concepts in the different disciplines. Entering a different field has imminent danger, with the

social sciences already making some fundamental mistakes via naïve candidate gene studies that have produced false positive findings that will likely fail to replicate. This concern was emphasized by Wachter (2008: 1503), who formulated this key challenge as follows: "the need to keep a high standard of quality control, as interdisciplinary researchers step out beyond their original base of expertise." This means not only borrowing the techniques or data from a particular field, but actively engaging with researchers in those fields and following training to ensure high quality standards. But this is an approach that begs for precious time and resources often unavailable for researchers and requires them (their employers and funders) to possess the ability to understand delayed gratification and long-term thinking, since investments into a new discipline may take years to bear fruit. In the current sphere of competitive research and focus on production, many researchers – and particularly young scientists – may be unable or reticent to make these kinds of long-term investments. Interdisciplinary training programmes and centres could be a valuable solution to train the next generation of truly interdisciplinary scholars.

1.4.4 Towards more complex models of gene x environment (GxE) interaction

We have shown that the environment can modify genetic influences (chapters 2 & 4), however, it is highly likely that there are multiple ways in which the socio-environment can moderate the genetic endowment of individuals, which remain unknown so far. There are two main ways in which genetic dispositions may influence human fertility. First, there can be a direct effect on physiological characteristics (e.g., fecundity, age at menarche, age at menopause). Second, biological predispositions may affect the processes of decision-making and life course planning, consciously or subconsciously (Kohler et al., 2006). Genes provide the potential for a trait, but environmental conditions determine whether that potential will be realized. The most promising future research will therefore be that which attempts to go beyond the examination of only direct effects of genetic and socio-environmental factors to explore their combined effect and interaction, which is likely greater than their independent effects. It would be interesting to explore how the socio-environment moderates gene expression. It is imaginable that triggers

such as stress might activate certain hormones and other bodily functions more prevalent in individuals with a particular genetic endowment which in turn could inhibit their fecundability. We also know that social compensation, in the form of social capital and strong networks can result in individuals being able to realize their initial fertility intentions (Balbo & Mills, 2011). PS, again, are a promising way to test for gene-environment interaction in statistical models.

1.4.5 Epigenetics

The growing interest and study of epigenetics is another relevant consideration for future research (Landecker & Panofsky 2013). This relates to the dynamic nature of genes and how their expression is influenced by DNA methylation. It is plausible that socio-environmental exposures might actually alter gene expression through changes in DNA methylation. For example, the study by Cameron et al. (2005) demonstrated that the maternal behaviour of rats (grooming, nursing) affected gene expression among offspring in the brain regions that control defensive and reproductive behaviours. Another recent study on rats also showed epigenetic silencing or the inhibition of DNA methylation as a mechanism underlying the neuroendocrine control of female puberty (Lomniczi et al. 2013). In humans, factors such as breast-feeding and parental interaction could be potentially interesting factors to study.

1.4.6 Sex differences

Almost identical results in heritability estimates for men and women as in Rodgers & Kohler (2001) might suggest that the same genes are important for male and female fertility. However, the study of Nisén et al. (2013) shows that genes predicting childlessness in women are associated with low education among women and high education among men. From a social sciences perspective, this is not implausible. We know a number of causal mechanisms which explain the association between educational level and fertility and may have different effects for both sexes (Balbo et al., 2013; Mills et al., 2011; Ní Bhrolcháin & Beaujouan, 2012;

Nisén & Myrskylä, 2014), for example opportunity costs of the transition to parenthood and wage penalty (Amuedo-Dorantes & Kimmel, 2005; Happel, Hill, & Low, 1984; Liefbroer & Corijn, 1999). Also biologically, human reproduction is complementary and therefore the genetic architecture of fertility might differ considerably between the sexes (Ellegren & Parsch, 2007).

In the recent past, several methods have been developed to approach this question for the first time. The GREML models we apply in this dissertation can be used to investigate the genetic correlation in fertility across sexes. Based on GWAS summary results, a new technique has been developed called LD score regression which is able to estimate the genetic correlation across traits and also across sexes on the same trait if the GWAS is conducted separately for women and men (Bulik-Sullivan et al., 2015). Both approaches have recently been used to investigate genetic difference across sexes for height and BMI, demonstrating strong genetic homogeneity for men and women. However, this remains to be investigated for fertility, which appears to be rather complementary: sex-biased and antagonistic gene-expression have been widely discussed and hypothesized in the literature (for review see e. g. Parsch & Ellegren, 2013).

1.4.7 A population paradox

Evolutionary predictions that there is a tendency to have children at a younger age over time seem to contradict widespread observations of fertility postponement at the population level in many European countries. Models from biologists suggest that the AFB decreases across generations (Byars et al., 2010; Kirk et al., 2001), while we observed a massive postponement during the second half of the past century (Mills et al., 2011). Potential explanations for this population paradox are, amongst others, gene environment interaction or additive environmental effects such as the educational expansion or the introduction of the pill overriding smaller effects of natural selection. Once more, this highlights that an integrative approach between biological and social science is necessary to answer questions of human fertility and evolution.

1.4.8 Integrating new data and methods from quantitative genetics and reproductive medicine

As mentioned at the outset of this introduction, a central reason for the lack of biodemographic research in the area of fertility has been a lack of data that combines good social science indicators with data that has biomarker and genetic data in large samples. The recent significant reduction in the costs of collecting, storing and processing this type of data will hopefully result in new sources for the future. Although considerable research has been done, many questions remain open. The fact that heritability is a population parameter and therefore can vary over time and space requires replication of heritability estimates in different societies of interest.

To break new ground in fertility research using genetic information, methodological advances such as the ACE-β model has been developed (see chapter 5). This approach combines insights from econometrics and behavioural genetics by integrating the MZ-fixed effects approach as a direct link between phenotypes into the structural equation model. It therefore potentially identifies a causal link between two traits additional to genetic and environmental endowment – particularly if a valid instrumental variable can be found (these and related approaches are detailed in Kohler et al., 2011). More importantly, in the realm of molecular genetics, actual genetic (GWAS) data is being collected and statistical tools further developed (Yang et al., 2010; Yang, Lee, Goddard, & Visscher, 2011). Using actual genetic information, it becomes possible to determine the heritability of one as well as the genetic correlation between two traits without family data – which is often problematic to collect and requires strong assumptions in the modelling (Conley, Rauscher, Dawes, Magnusson, & Siegal, 2013). This likewise raises the level of flexibility in testing gene environment interactions and sex differences.

Additional techniques have also been developed within quantitative genetics to try to explain more of the variance predicted in behavioural genetics models or in other words to address the 'missing heritability' problem. Several methods combine many SNPs to aggregate the collective effect and achieve a higher level of predictive power such as polygenic scores (Purcell et al. 2009; Mandemakers et al. 2014) or GREML implemented in genome-wide complex trait analysis software (GCTA, see chapters 3 & 4) (Yang et al. 2010). Finally, other promising applications include using genes as instrumental variables to go beyond establishing association and make

causal inferences. Endogeneity problems are often difficult to disentangle, with these approaches offering at least a partial solution (for a critical discussion see Conley, 2009; Taylor, Ware, & Gage, 2015).

A central conclusion from this thesis is that biological and genetic factors are relevant in explaining and predicting fertility traits. The socio-environment, which demographers and social scientists have studied until now, however, still appears to account for the major part of the observed variance. Studying the interplay between genes and the environment, new data sources and integration of new methods will be central to more effectively understanding and predicting fertility trends. Findings, such as the common genetic base of the first attempt to have a child and NEB furthermore suggest that genes not only underly physical traits important for fertility but also psychological ones. In the coming years we anticipate that the field and interest will grow as we become able to isolate which genetic variants are important for fertility, understand their biological functions and how they interact with the socio-environment.