BEHAVIOR GENETIC MODELING OF HUMAN FERTILITY: FINDINGS FROM A CONTEMPORARY DANISH TWIN STUDY*

JOSEPH LEE RODGERS, HANS-PETER KOHLER, KIRSTEN OHM KYVIK, AND KAARE CHRISTENSEN

Behavior genetic designs and analyses can be used to address issues of central importance to demography. We use this methodology to document genetic influence on human fertility. Our data come from Danish twin pairs born from 1953 to 1959, measured on age at first attempt to get pregnant (FirstTry) and number of children (NumCh). Behavior genetic models were fitted using structural equation modeling and DF analysis. A consistent medium-level additive genetic influence was found for NumCh, equal across genders; a stronger genetic influence was identified for FirstTry, greater for females than for males. A bivariate analysis indicated significant shared genetic variance between NumCh and FirstTry.

The research presented here combines a design and an analytic approach that are not typically found in demographic literature with an outcome variable that is of central interest to demographers. To study fertility, we present a set of behavior genetic analyses based on a twin design. In this paper we address two basic questions. First, are there genetic influences on fertility outcomes? Second, if so, what are the theoretical origins of those influences?

Behavior genetic designs are not often used in demography. During the past decade, however, increasing attention has been given to integrating biological explanations with the type of social models used in demography and sociology, particularly in relation to human reproduction (e.g., Adams et al. 1990; Rossi 1994; Udry 1995; Wood 1994). It is unfortunate that behavior genetic methods are not standard tools in demography, because they provide a powerful set of designs and analyses to address questions of interest to demographers. Further, the term *behavior genetics* is a misnomer, to some extent: Plomin and Rende (1991) commented, "The power of behavioral genetics lies in its ability to consider nurture as well as nature—that is, environmental as well as genetic sources of individual differences in behavior" (p. 162).

Even if a researcher's primary goal is to study social/environmental influences on behavior, the ability to control for genetic confounds is often critical in this assessment. Scarr and Grajek (1982) took an extreme position on this issue, suggesting that "[p]assive genotype-environment correlations arise in biologically related families and render all of the research literature on 'socialization' uninterpretable" (p. 374). Rodgers, Rowe, and Li (1994) presented an analytic approach that combines genetic and environmental sources of influence, allowing for the control of either in the evaluation of the other.

The primary purpose of this paper is to present a behavior genetic analysis of human fertility and to provide substantive interpretations of the results. As a secondary and concomitant goal, however, the paper can serve as a didactic tool to introduce demographers who are unfamiliar with behavior genetic methods to this class of techniques. We do not attempt exhaustive coverage of the models and assumptions that underlie behavior genetics, but we provide broad descriptions and make references to the literature that supports those descriptions. Further, the study of genetic influences on human fertility naturally leads to theoretical treatment of human evolution and Fisher's fundamental theorem of natural selection (FTNS). Again, we cannot provide exhaustive coverage of the FTNS, but we provide summary treatment along with references to the biological literature, which interested readers may consult.

Obviously this paper contains both methods and theory that are not typically found in *Demography*. This type of work, however, can have a substantial impact on demographic research and thinking. We provide one motivational example. Using a larger and overlapping part of the Danish twin data that we analyze here, Kohler, Rodgers, and Christensen (1999) estimated a time series of heritability coefficients associated with human fertility in cohorts born from 1870 to 1910 and from 1953 to 1963. This historical analysis showed that the role of genetic influences in accounting for fertility outcomes changed substantially over time, ranging from zero for some periods to a moderate level of around .40. (That is, this finding suggested that around 40% of the observed variance in family size was associated with genetic variance on fertility.) In addition, the role of the family environment varied in a systematic and interpretable fashion. It is a matter of broad interest that heritability coef-

*Joseph Lee Rodgers, Department of Psychology, University of Oklahoma, Norman, OK 73019; E-mail: jrodgers@ou.edu. Hans-Peter Kohler, Max Planck Institute for Demographic Research; Kirsten Ohm Kyvik, Section for Epidemiology, Institute for Public Health, and The Danish Twin Registry, University of Southern Denmark; Kaare Christensen, Section for Epidemiology, Institute for Public Health, and the Danish Center for Demographic Research, University of Southern Denmark. The authors acknowledge contributions from reviewers and the editors that substantially improved this paper. The first author was supported by NIH Grant RO1-HD2-1973 during the course of this research. Research support and space were also graciously provided to him by the Danish Center for Demographic Research at Odense University and by the Max Planck Institute for Demographic Research in Rostock, Germany during a sabbatical leave in 1998. The second author acknowledges the support provided by the Max Planck Institute for Demographic Research. The third and fourth authors were supported by Sygekassernes Helsefond (Grant 11/209-93). Administrative support from Jim Vaupel and Hans Christian Johansen was particularly valuable in stimulating the collaboration that led to this paper.

ficients of fertility should change over time, but the relation of those changes to the demographic transition is especially notable for demographic analysis.

Two demographic transitions occurred in Denmark: the first for cohorts born during the 1880s, the second for cohorts born during the 1950s (van de Kaa 1987). Heritabilities increased dramatically at the beginning of each transition. This finding helps us to understand the fertility changes underlying the demographic transition in Denmark (and potentially in other countries as well). The analysis presented here focuses on a contemporary part of the historical pattern set forth in Kohler et al. (1999). In the discussion section we further develop the relation between genetic influences on human fertility and the demographic transition, and we provide additional description and interpretation of the Kohler et al. (1999) results.

THE GENETICS OF FERTILITY AND FISHER'S FTNS

One of many threads running through the recent behavior genetic literature is identification of genetic influences on unexpected domains of human behavior (several of which are particularly interesting to demographers). Examples include research showing genetic influence on divorce (McGue and Lykken 1992), television viewing (Plomin et al. 1990), happiness (Lykken and Tellegen 1996), religion (Waller et al. 1990), and adolescent sexual behavior (Dunne et al. 1997; Rodgers, Rowe, and Buster 1999). Yet the most important repertoire of behavior for the survival of our species—human fertility—has seldom been studied. In fact, a review of the psychological and behavior genetic literature revealed only one study (Mealey and Segal 1993) that applied modern behavior genetic modeling to fertility outcomes. A few articles have treated genetic influences on human reproduction from a distance. For example, Perusse et al. (1994) considered the heritability of child-rearing practice and parental bonding. The treatment by Dunne et al. (1997) and Rodgers et al. (1999) of age at first intercourse accounted for reproductive behavior, although neither study addressed how adolescent sexual behavior ultimately translates into fertility outcomes.

Superficially, considering genetic influences on fertility behavior has the same nonsensical appearance as the old joke about whether celibacy can be passed from parents to children. If one or more genes coding directly for large families existed in humans, we would be hard pressed to explain many features of human behavior, including modern contraceptive practices, postponement of childbearing to pursue educational and occupational goals, declining family size following the fertility transition, and increasing levels of voluntary childlessness. Although asking questions about genetic influence is much more subtle than such a deterministic model would imply, researchers have shown little interest in including genetic influences even as small parts of models of reproduction and fertility. (For comments on this position, see Newcomer 1994; Udry and Campbell 1994.)

Sir Ronald Fisher is at least partially to blame. Fisher's fundamental theorem of natural selection (FTNS; Fisher 1930) showed that for traits and behaviors that make strong contributions to reproductive fitness, the additive genetic variance will disappear in the long run. This result implies that such traits will have heritabilities (which measure the proportion of phenotypic variance associated with genetic variance) of approximately zero, and therefore no genetic etiology. Thus, on its surface, the FTNS implies that it would be useless to search for genetic influences on sexual behavior, much less on fertility outcomes. Ironically, however, in the same book in which Fisher presented the FTNS, he estimated the heritability of completed family size in a sample of British aristocrats, and found $h^2 = .40$. Nevertheless, it appears that social scientists and behavioral geneticists have taken his theorem more seriously than his empirical results.

This is not true of other fields, however, particularly biology. At least four different biological journals have published recent expositions on Fisher's FTNS (see Edwards 1994; Ewens 1989; Frank and Slatkin 1992; Lessard 1997; also see Price 1972). These and similar articles justify an effort by social scientists to seek and explain patterns of genetic influence on reproductive behavior. Was Fisher wrong? Certainly not; in fact, current literature suggests that Fisher's theorem is still ahead of our time (e.g., Edwards 1994), but it applies in the context of a certain set of assumptions. Further, it has been consistently misinterpreted, even by specialists. We present a short and simplified explanation of why Fisher's theorem can be right, while we can still fruitfully study genetic influences on fertility and reproduction.

Fisher argued that a population's average reproductive fitness increases because of natural selection, but decreases back into an equilibrium because of mutations and environmental changes (e.g., Burt 1995). His FTNS showed mathematically that the genetic variance in fitness traits should approach zero in the limit. As Houle (1992:197) explains, however, genetic variance and therefore heritability will decline to zero only *"in the absence of perturbing forces*" (our emphasis). Many believed that Fisher was stating a result for both sides of the equilibrium, although recent reinterpretations show that he intended the theorem to apply only to the natural selection side (Frank and Slatkin 1992). Thus, if both mutations and environmental changes can introduce genetic variance back into the process, the search for genetic influences on fertility-related behaviors should be an ongoing activity. (In fact, a number of other sources also can introduce genetic variance back into the system, including frequency-dependent selection, heterozygote advantage, and sexual antagonism, although these processes are probably not as important as mutation. For elaboration see Hughes and Burleson 2000; Rodgers et al. forthcoming.)

Have perturbing forces been operating on human reproductive behavior? Perturbing forces are environmental changes that interact with the genotype and its phenotypic expressions. One example, discussed by Fisher himself, is contraceptive innovation (Fisher 1930). Other examples would be changes in societal norms for sexual attraction, changes in marriage and cohabitation (and thus mating) patterns, changes in desires and norms regarding family size,

changes in availability of induced abortion, and efficacious treatment of human infertility. Obviously the modern world is full of "perturbing forces." (For a more technical specification of such forces, see Mayo, Burger, and Leach 1990.) Thus, without even invoking the issue of genetic mutation (which in fact may be even a more powerful source of genetic variation than these other influences), we have abundant motivation to search for, measure, and explain patterns of genetic influence on human reproduction and fertility.

As an added motivation, especially for social scientists, accounting for genetic influence can more fully inform social and environmental models of reproduction and fertility. Miller has emphasized the important role in fertility outcomes played by fertility motivation, which in turn is influenced by social/personality factors (e.g., Miller 1992) and biological/genetic influences (e.g., Miller et al. 1999). Some scholars have argued that the extensive social theorizing about human fertility that has occurred over the past several decades among psychologists, sociologists, demographers, economists, and public health experts is fundamentally flawed unless the social models are informed as well by biological/genetic components (e.g., Udry 1995).

The literature offers little empirical encouragement to seek genetic influences on human fertility. Fisher's (1930) heritability estimate of $h^2 = .40$ was criticized by Williams and Williams (1974) as being an artifact of secular fertility trends. Imaizumi, Nei, and Furusho (1970) found nonsignificant heritabilities in Japanese data from 1881 to 1930 in analyses of father-child ($h^2 = -.02$) and mother-child ($h^2 = .12$) correlations. Mealey and Segal (1993) used data from U.S. twins raised apart, and found a monozygotic twin (MZ) correlation of .06 (implying $h^2 = .06$; $n = 32$ pairs) and a dizygotic twin (DZ) correlation of $r = .10$ (implying $h^2 = .20$; $n = 23$ pairs); neither was significant. These findings would lead one to believe that heritability in human fertility is indeed around zero. These studies, however, were based on relatively small data sets with low external validity; only Mealey and Segal used modern behavior genetic methods in their analysis.

METHODS

Data

Our data come from the Danish Twin Registry for a cohort born between 1953 and 1959. The sample included all surviving twins born in Denmark during this period who responded to a set of survey questions administered in 1994, when these twins ranged in age from 35 to 41. (For recruitment details, see Christensen et al. 1998 or Kyvik, Green, and Beck-Nielsen 1995.) Opposite-sex DZ twin pairs were excluded because there is no MZ comparison group. The sample contained 3,240 twins (1,620 pairs): 1,226 MZ twins (613 pairs) and 2,014 DZ twins (1,007 pairs). There were 650 MZ and 1,046 DZ male-male twin pairs, and 576 MZ and 968 DZ female-female twin pairs. Kohler et al. (1999) and Christensen et al. (1998) investigated whether the Danish twins are similar to the overall population in their fertility behavior. They found minor differences at certain points in Danish history, but in general the twins' fertility behavior can be considered as very similar to that of the Danish population.

Twins provide researchers with a powerful design to investigate the role of both genetic and environmental influences in accounting for individual differences. MZ twins share 100% of their genes; DZ twins share, on average, 50% of their genes. Again, Fisher was the first to take advantage of this genetic structure to develop mathematical theory supporting research from a biometrical perspective (e.g., Fisher 1918). Modern quantitative genetic models (e.g., Falconer 1981) show how correlations defined between twin pairs can be used to estimate the proportion of variance in an outcome that is directly attributable to genes, to the shared environment (i.e., the family environment that siblings share, which acts to make them similar to one another), and to the nonshared environment (i.e., within-family influences that act to make siblings different from one another) and measurement error.

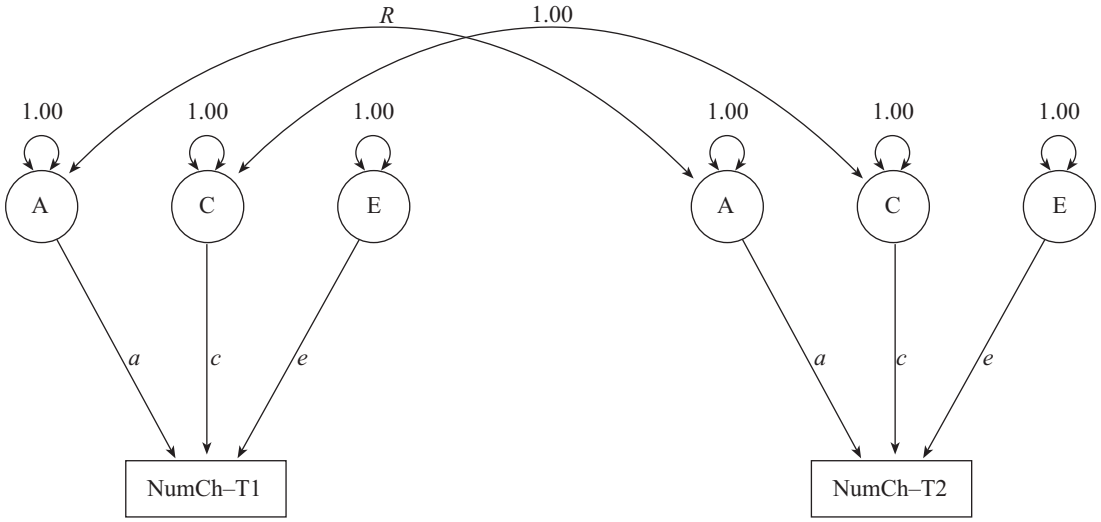
Behavior genetic terminology (see, e.g., Neale and Cardon 1992) often refers to the genetic portion of such models as A (for "additive genetic variance"), the shared-environment portion as C (for "common environment"), and the nonshared-environment/measurement-error portion as E (for "error"). Thus an ACE model contains all three sources of variability, while an AE model has dropped the shared environmental component. Twin designs in which MZ and DZ twin correlations are compared with each other have certainly been the most popular type of biometrical/behavior genetic design, although many other types of kinship designs can be used in this type of research as well (e.g., adoption designs, sibling versus half-sibling designs, parent-child designs, or designs using identical twins raised apart). A highly accessible treatment of behavior genetic designs and analyses can be found in Plomin (1990); Plomin, DeFries, and McClearn (1990) have written a textbook providing more in-depth treatment.

The Analytic Models

Two different modern approaches are available for estimating the parameters in behavior genetic models: structural equations modeling (SEM) and a regression method called DF analysis. SEM approaches use maximum likelihood as the fitting criterion; they require not only a structural model explaining the relation between the constructs of the model but also a measurement model explaining the relation between the constructs and the variables used to measure the constructs. Mx (Neale et al. 1999) is a statistical software package that implements SEM methodology specifically to estimate behavior genetic models. Alternatively, other SEM estimation procedures such as those in LISREL or EQS can be used.

An example of an ACE model is presented in Figure 1. The boxes correspond to the fertility variables "Number of Children, Twin 1" (NumCh–T1) and "Number of Children, Twin 2" (NumCh–T2). These are the measured variables. The latent variables are contained in the circles, and corre-

FIGURE 1. ACE MODEL FOR NUMBER OF CHILDREN (NumCh); A = GENETIC INFLUENCE, C = SHARED ENVIRONMENTAL INFLUENCE, AND E = NONSHARED ENVIRONMENTAL INFLUENCE/MEASUREMENT ERROR



spond to additive genetic, common environmental, and nonshared environment/error sources of variance. The *R* coefficient linking the path between the two twins' genetic component measures the degree of genetic relatedness: $R = 1.0$ for monozygotic (MZ) twins, $R = .5$ for dizygotic (DZ) twins. The small letters *a*, *c*, and *e* are structural coefficients. If they are standardized (or if correlations rather than covariances are fitted), then a^2 estimates heritability, h^2, and c^2 provides a direct estimate of shared environmental variance (see Neale and Cardon 1992). Mx recently has implemented a graphical-user interface (GUI) procedure in which a model such as that in Figure 1 can be defined graphically and linked to a file containing covariances or correlations to estimate the parameters in the model.

The second estimation strategy, DF analysis, was developed by DeFries and Fulker (1985); they devised a method that uses least-squares regression, a less technical computational approach than that employed in SEM. Historically, most behavior genetic parameters reflecting genetic or environmental influence were obtained by comparing kinship correlations across pairs of kinship categories (e.g., comparing MZ with DZ twin correlations, or adoptive-sibling with full-sibling correlations), and using those correlations in formulas that provide estimation of heritability, h^2, and shared environmental variance, c^2. DF analysis provides a regression model that simultaneously estimates h^2, c^2, and nonshared environmental/error parameters within the same model (i.e., a complete ACE model). Although SEM provides more flexible and more powerful methods, DF analysis now is almost as broadly applicable as SEM, and can be implemented easily from standard statistical packages with least-square regression procedures. We use both approaches in fitting fertility data from the Danish Twin Registry.

The analytic basis for DF analyses presented here comes from the following set of regression models:

$$\text{FM1} = b0 + b1 \times \text{FM2} + b2 \times R + b3(\text{FM2} \times R) + e \quad (1)$$

$$\text{FM1} = b0 + b2 \times R + b3(\text{FM2} \times R) + e \quad (2)$$

$$\text{FM1} = b0 + b1 \times \text{FM2} + b2 \times R + e \quad (3)$$

$$\text{FM1} = b0 + b2 \times R + b3(\text{FM2} \times R) + b4(\text{FM2} \times D) + e \quad (4)$$

$$\text{FMB1} = b0 + b1 \times \text{FMA2} + b2 \times R + b3(\text{FMA2} \times R) + e \quad (5)$$

$$\text{FMB2} = b0 + b1 \times \text{FMA1} + b2 \times R + b3(\text{FMA1} \times R) + e \quad (6)$$

$$\text{FM1} = b0 + b1 \times \text{FM2} + b2 \times \text{R} + b3(\text{FM2} \times \text{R}) + b4 \times \text{GENDER} + b5(\text{GENDER} \times \text{FM2}) + b6(\text{GENDER} \times R) + b7(\text{GENDER} \times \text{FM2} \times R) + e. \quad (7)$$

Eq. (1) is the basic DF analysis model (an ACE model). FM1 and FM2 are fertility measures (or other outcomes in other types of research) from Twin 1 and Twin 2 (Kin 1 and Kin 2 in broader settings), *R* is the coefficient of genetic relatedness ($R = 1.0$ for MZ twins, $R = .5$ for DZ twins), *e* is the error/residual term, and the *b*s are least-squares regression

coefficients. Given the assumptions of the additive genetic model (equal environments for MZ and DZ twins, trivial assortative mating, and additivity; see Falconer 1981 for extensive treatment of the quantitative genetic model and its assumptions), $b3$ provides an unbiased estimate of the proportion of phenotypic variance associated with genetic variance (heritability, or h^2); $b1$ provides an unbiased estimate of the proportion of phenotypic variance associated with shared environmental variance, or c^2 (see DeFries and Fulker 1985 or Rodgers and McGue 1994). In cases of ambiguity about which twin should contribute his or her score to FM1 and which to FM2, the data are double-entered so that an observation exists for each individual, or twice the number of kinship pairs. Significance tests can be conducted conservatively, using the sample size of twin pairs and adjusted standard errors; all of our significance tests reflect this adjustment.

If either $b1$ or $b3$ is approximately zero in Eq. (1), a more parsimonious model can be reestimated, which omits this component. For example, if Eq. (1) estimates $b1 = c^2 = .02$ and $b3 = h^2 = .40$, a genetic-only model (an AE model) can be reestimated using Eq. (2). Eq. (3) provides the equivalent model to estimate a shared-environment-only model (a CE model). The legitimacy of this refitting is based on the same grounds as stepwise regression, in which useless variables are dropped from a regression equation and the equation is refitted in search of a best-fitting model.

If $b1$ in Eq. (1) is estimated to be a negative value (a violation of the additive genetic model), then the model can be reestimated to test for dominance, one of the two types of genetic nonadditivity, with an ADE model (where D refers to the dominance component), as shown by Waller (1994). The other type of genetic nonadditivity—epistasis—is very difficult to test for, and usually is not treated in such modeling efforts (see Neale and Cardon 1992:153). In Eq. (4), D is the dominance coefficient ($D = 1.0$ for MZ twins, $D = .25$ for DZ twins), $b4$ estimates the dominance effect d^2, $b3$ estimates h^2 (narrow-sense heritability), and $b3 + b4$ estimates broad-sense heritability. Narrow-sense heritability refers to additive genetic variance; broad-sense heritability to both additive variance and interactive variance (including both dominance and epistasis). In a twin design, the ACDE model cannot be estimated because of the confounding between pairs of the A, C, and D components.

The DF analysis represented in Eqs. (5) and (6) is a new application of the DF approach. We use two qualitatively different measures of fertility; one precedes the other temporally. We wish to learn whether the environmental and/or genetic variance in the earlier fertility behavior accounts for variance in the later behavior—a cross-variable DF analysis. In Eq. (5), the first fertility variable is FMA; the second is FMB. (As before, we use numbers to denote Twin 1 or Twin 2.) The genetic variance in FMB that is explained by FMA is estimated by $b3$, and the shared environmental variance by $b1$. Eq. (5) must yield results identical to those of the other cross-variable model represented in Eq. (6). Because the DF analysis procedure requires FMA and FMB to be measured on the same scale (see Rodgers and McGue 1994), FMA and FMB (which may have been measured on different scales) must be standardized in the sample to have equal means and standard deviations. This type of analysis is implemented easily with SEM methods, and is typically called a bivariate analysis. More flexible and broader bivariate SEM analyses can be implemented; one of these will be presented later in this paper.

Because males and females play different roles in the reproductive process, we fitted our models separately by gender. Tests for gender differences can be made in either DF analysis or SEM approaches. In DF analysis, this can be achieved by adding a gender dummy variable and several gender interactions to the basic DF analysis model, as in Eq. (7) (LaBuda and DeFries 1990). Within Eq. (7), $b5$ provides a direct test of whether c^2 differs by gender, and $b7$ provides a test of whether h^2 differs by gender.

DF analysis originally was developed to handle a specific problem associated with extreme forms of a measured variable (DeFries and Fulker 1985). The genetic and environmental variance related to the normal range of a trait may be different from the variance related to extremes. For example, Rodgers et al. (1999) investigated whether heritability of age at first intercourse defined over the whole distribution was the same as the heritability in a sample with early first intercourse, and in a second sample with late first intercourse.

The same type of question arises in relation to fertility: Do genetic/environmental influences on the whole range of fertility outcomes match those from high- or low-fertility individuals? To estimate h^2, c^2, and d^2 associated with an extreme trait, proband and co-twin links can be specified, and the models in Eqs. (1) to (7) can be fitted to the (nondouble-entered) data (e.g., LaBuda, DeFries, and Fulker 1986; Rodgers, Rowe, and Buster 1999). The probands represent individuals who have been selected because of their status as extreme on some particular trait (e.g., reading disability, late age at first intercourse, extreme delinquency). In these models, the FM2 (independent) variable corresponds to the proband, and the FM1 (dependent) variable to the proband's co-twin. In fact, the first application of DF analysis was to model genetic influences on reading disability. We use this type of analysis to evaluate whether genetic influences for low or high fertility appear to have the same etiology as those for the general fertility distribution.

In a final analysis, we compute coefficients of genetic variation to supplement the information provided by heritability estimates. Heritabilities are proportions, and thus "wash out" the information about overall phenotypic or genetic variance. Coefficients of genetic variation reflect the overall variance, standardized against the phenotypic mean. We compute these from heritability estimates obtained from DF analysis, along with the sample means and standard deviations of our measures, using the following formula:

$$\begin{aligned} CV_a &= 100\,\sqrt{V_a}\,/\,M_p = 100\,\sqrt{V_p}\,\sqrt{h^2}\,/\,M_p \\ &= 100\,S_p(\sqrt{h^2})\,/\,M_p, \end{aligned} \tag{8}$$

where V_a is additive genetic variance, V_p is phenotypic variance, M_p and S_p are the phenotypic trait mean and standard

deviation, and h^2 is the heritability coefficient. $S_p(\sqrt{h^2}) / M_p$ is obtainable from the observed phenotypic mean, standard deviation, and estimated heritability of the fertility measure. Houle (1992) and Burt (1995) emphasized the importance of using direct estimates of genetic variance for comparison across domains, because heritabilities are influenced by the amount of phenotypic variance that is to be explained. Houle noted that heritabilities usually are lower, the closer a measure is to a pure fitness measure (e.g., completed fertility), whereas coefficients of variation usually have the opposite relation to fitness. Houle analyzed 842 estimates of genetic variance of fitness traits, observed consistently high coefficients of variation, and attributed those to the many different sources of environmental influence (i.e., perturbing forces).

In summary, in our analytic approach we fitted fertility kinship correlations from the Danish Twin Registry, using both DF analysis and SEM approaches. Because it is less well known but also easier to implement (and therefore more accessible), we present the DF analysis first. Then we present equivalent results using SEM. The DF analysis provides particular flexibility for addressing the extremes of the fertility distribution; the SEM analysis provides broader treatment of the bivariate relation between different fertility measures. We apply both capabilities to these data.

Measures

The two fertility measures we use are number of children (NumCh), a traditional measure of completed fertility, and age of first attempt to get pregnant (FirstTry). Truncation in these variables is a matter of some concern because they can be incomplete for 35- to 41-year-olds. This is only a minor problem in the Danish twin data, however: the distribution of FirstTry shows that only 1.2% of respondents had a FirstTry equal to 35 or older. National Danish fertility data (Danmarks Statistics 1997) suggest that in 1994, approximately 11% of births occurred at age 35 or after, and about 3.5% of first births occurred after age 35. Thus truncation on NumCh is well below 11% (because 35 is the age of the youngest twins in this sample). This truncation issue is treated more formally in Kohler and Rodgers (1999), who used bivariate tobit models.

FirstTry is a new and valuable fertility measure. It was obtained from a 1994 survey sent to all respondents in the Danish Twin Registry, from a question asking "What was the year of your first attempt to have a child?" We know of no previous use of this measure, although similar measures can be found in the literature. (See, for example, Martin and Treloar 1991, who used a measure of age at first pregnancy.) Completed fertility—measured as number of children—accounts for many different potential processes. Influences on NumCh include fertility desires, desires of partner(s), contraceptive failure, infertility, family size norms, and random factors. Thus, completed family size measures a combination of attitudinal, behavioral, and random characteristics of both the individual and others. FirstTry, however, is a behavioral indicator that is not contaminated so strongly by contraceptive failure, infertility, or other random factors. It acts as a precursor to fertility for those individuals who plan their pregnancies successfully. We view this new measure as a potentially valuable (though imperfect) measure of *fertility motivation*, especially in its application to the Danish culture. Because it has not previously been used extensively, we discuss its strengths and weaknesses in some detail.

In Denmark we believe that this measure may behave somewhat differently than in many other cultures. The Danish culture is characterized by relatively early first intercourse (e.g., Rowe and Rodgers 1991), by early and almost universal contraceptive use, and by a great deal of out-of-wedlock childbearing. For example, Helwig-Larsen, Knudsen, and Petersson (1998) reported that in the late 1980s, about 35% of adult Danes lived in consensual unions with children (also see Lopez and Cliquet 1984). In this type of culture, there is less contraceptive failure than in the United States and greater control over fertility. Further, fertility behavior is separated somewhat from marital status. These observations imply that the age at which an individual reports first trying to get pregnant, in the Danish culture, should reflect some combination of fertility motivation and other activities—such as education—that compete with childbearing.

The measure FirstTry raises at least two concerns. The first is the potential for retrospective bias. To answer this survey question, respondents in 1994 were asked to recall an event that had occurred, for some, about 20 years earlier (and for others, only shortly before 1994). Although one may question whether such a memory is reliable, there are reasons to believe that it is. Bradburn, Rips, and Shevell (1987), discussing autobiographical memory, suggested that certain events are remembered better than others; those associated with fertility appear to be markers in defining a person's memory of his or her own past. Previously we argued for the reliability and validity of measures of age at first intercourse (e.g., Rodgers, Harris, and Vickers 1992), partially on the grounds that individuals often organize their memories in relation to events that are socially defined as transitional markers. The first attempt to become pregnant may be such an event (although we hasten to encourage additional research on this question). Both early and late sexual debut may be personally embarrassing, and therefore can evoke incorrect or evasive answers. We suspect, however, that this type of bias occurs more often in the "early" direction for FirstTry; a late age for first attempt to become pregnant can be justified easily in relation to education or marriage. Thus we conclude by noting the potential for retrospective bias, but we suspect that most respondents remember fairly well the year in which they first tried to get pregnant.

The second weakness of FirstTry is that for some respondents it precedes first birth, while for others (whose first pregnancy is accidental) it follows first birth. In the Danish twin data that we used, fewer than 5% of the respondents indicated that they had tried to get pregnant but had never had a child, and barely more than 5% reported that they had never tried to get pregnant, but had had a child. This point, however, still leaves open the question of timing. Age at first birth has only recently become available for the Danish twins, and an exten-

sive analysis of that new variable is in progress. Some preliminary results suggest that 27 females and 52 males report a negative difference between FirstTry and year of first birth (3.4% of the 2,353 twins for whom neither variable was missing). For most of these 79 individuals, FirstTry precedes the *second* birth by a short time, suggesting that the negative differences usually represented unintended first births, followed by desired second births. These numbers suggest that, at least in the Danish culture, FirstTry typically precedes first birth. In addition, however, we argue that FirstTry is a useful measure *regardless of its relation to timing of first birth*. If FirstTry is considered to be a measure of fertility motivation, then an unwanted first birth does not damage this interpretation; indeed, FirstTry should not respond to unwanted births if it is measuring fertility motivation.

In the discussion section, we further develop the idea that FirstTry is a useful (though imperfect) measure of fertility motivation, especially in the presence of childbearing and marital circumstances such as those in contemporary Denmark. In general, however, we consider this measure to be a reliable fertility precursor—a measure that provides information about the potentiality of fertility behavior. In this sense it is parallel to measures of fertility expectations, fertility desires, and fertility intentions. The primary difference between our measure and these other fertility precursors is that FirstTry is a behavioral measure, like age at marriage. Unlike age at marriage, however, it is a direct and unambiguous indicator of intention to have children.

We distinguish theoretically between two domains of influence on ultimate fertility. The first domain is purely biological, reflected by the idea of fecundity. The definition of fecundity—an individual's physiological capacity to reproduce (e.g., Shryock and Siegel 1976)—does not leave much latitude for volitional behavior. The second domain is behavioral/volitional; it reflects an individual's explicit and conscious effort to become pregnant. We believe that our measure FirstTry is an excellent indicator of this type of domain. We note that biological processes can influence this second domain as well: a person's volitional behavior may reflect many biological processes. For example, an individual might begin trying to get pregnant—a volitional activity—because of biological (possibly genetic) motivation to do so.

The role of volition seems to distinguish most clearly between these two domains. If indeed there are genetic influences on fertility outcomes, they could arise from either (or both) of these conceptual domains. Social scientists naturally would take a particular interest in the volitional variable; that is the one we investigate here. We will evaluate—through cross-variable DF analysis and through bivariate SEM—whether the "age at first attempt to get pregnant" variable accounts for genetic variance in NumCh, our fertility outcome measure.

To summarize, both of our measures—NumCh and FirstTry—are behavioral. One is the ultimate indicator of reproductive success; the other is (usually) a fertility precursor that behaves somewhat differently in Denmark than in many other cultures. In the Danish twin data, the overall sample correlation between NumCh and FirstTry was $r = -.33$. Thus, although these measures obviously overlap, they empirically account for rather different features of fertility behavior.

RESULTS

In Table 1 we present basic descriptive statistics for FirstTry and NumCh for the Danish twin sample. Christensen et al. (1998) document that these twins are very similar to the rest of the Danish population in their fertility behavior.

DF Analysis

Our first set of DF analyses was based on Eqs. (1) to (4). In Table 2 we summarize the results of this analysis for our two dependent variables. In each case, the genetic-only (AE) model (Eq. (2)) provided the best fit to the data (assessed by using F-tests to compare R^2 goodness-of-fit measures within these nested models; e.g., Maxwell and Delaney 1990:255). Estimated heritabilities were highly significant. Values of c^2 were consistently close to zero, and no dominance effects were detected. Heritabilities were higher for FirstTry than for NumCh. In comparisons across genders, h^2 values were virtually identical for males and females for NumCh, but Eq. (7) showed that females had significantly higher h^2 values for FirstTry ($t = 2.24$, $p < .01$).

The sample sizes in Table 2 are notably smaller for FirstTry than for NumCh because a number of respondents were missing on FirstTry: they had never attempted to get pregnant. Omission of this group from the analysis can contribute selection bias. In response, we reassigned missing values on FirstTry to be the respondent's age at the time of the interview in 1994. The third set of estimates in Table 2 corresponds to this analysis. The pattern of results was similar to that in which the actual responses were used; in fact, they were close enough to cause little apparent concern about this potential source of bias. In the remaining analyses, we con-

TABLE 1. DESCRIPTIVE STATISTICS, AGE AT FIRST ATTEMPT TO GET PREGNANT (FirstTry) AND NUMBER OF CHILDREN (NumCh)

	FirstTry	NumCh
Means		
Overall	26.3	1.54
Female MZ	25.3	1.65
Female DZ	25.5	1.58
Male MZ	27.0	1.54
Male DZ	27.4	1.45
Other Statistics, Overall Distribution		
Median	26.0	2.0
Standard deviation	4.4	1.08
Range	[16,40]	[0,6]
Bottom %	21.6% at or below 22	23.2% at 0
Top %	18.6% at or above 30	15.9% at or above 3

TABLE 2. FITS OF DF ANALYSIS MODELS (EQS. (1–4)) TO DANISH TWIN DATA FOR 35- TO 41-YEAR-OLD TWINS

					ACE Model			AE Model		ADE Model		
	r_{DZ}	N	r_{MZ}	N	h^2	c^2	N	h^2	N	h^2	d^2	N
Number of Children												
Male-male pairs	.13	1,032	.29	646	.32	–.03	1,678	.28[a]	1,678	.22	.07	1,678
Female-female pairs	.13	964	.31	576	.34	–.04	1,540	.29[a]	1,540	.23	.07	1,540
Age at First Pregnancy Try												
Male-male pairs	.19	596	.34	414	.31	.03	1,010	.35[a]	1,010	.41	–.07	1,010
Female-female pairs	.28	652	.51	400	.46	.05	1,052	.53[a]	1,052	.62	–.10	1,052
Age at First Pregnancy Try (Imputed Older Ages)												
Male-male pairs	.06	1,046	.23	650	.34	–.11	1,696	.20[a]	1,696	.22	.07	1,696
Female-female pairs	.19	968	.39	576	.41	–.02	1,544	.39[a]	1,544	.23	.07	1,544
Cross-Variable Analysis (NumCh From FirstTry)												
Male-male pairs	.04	592	.11	410	.15	–.04	1,002	.10[a]	1,002	.08	.03	1,002
Female-female pairs	.12	650	.16	400	.08	.08	1,050	.18[a]	1,050	.32	–.16	1,002

Notes: Sample sizes are numbers of individual twins. We divided these sample sizes by 2, measuring the number of twin pairs, to conduct significance tests. The ACE model contains components accounting for additive genetic variance, common environmental variance, and nonshared environmental variance/measurement error; the AE model contains additive genetic variance and nonshared environmental variance/measurement error; the ADE model contains additive genetic variance, dominance genetic variance, and nonshared environmental variance/measurement error.

[a]Best-fitting models.

tinue to use the FirstTry scores actually indicated. (For a more sophisticated estimation procedure to control for right-truncation in this type of setting, see Kohler and Rodgers 1999.)

Next we estimated the cross-variable model (Eqs. (5) and (6)), predicting NumCh for one member of the twin pair and FirstTry for the other. For this analysis, we could use only observations for which values were not missing for any of the four variables AgeFrst1, AgeFrst2, NumCh1, and NumCh2; total N = 2,052 (325 DZ female pairs, 200 MZ female pairs, 296 DZ male pairs, 205 MZ male pairs). The fourth set of estimates in Table 2 shows that for males, the best-fitting model was the genetic-only AE model, with a significant h^2 = .10. For females, the genetic-only AE model and the shared-environment-only CE model (not shown) were the best-fitting models, and they predicted equally well. The genetic-only model had a significant h^2 = .18; the shared-environment-only model had a significant c^2 = .13. Neither gender showed significant dominance effects. Although other factors as well obviously contributed to the genetic portion of NumCh, the age at which pregnancy was first attempted played a notable role through the genetic influences on FirstTry.

Then we estimated models for the tails of the fertility distribution using Eqs. (1) to (4), applied to the selected parts of the distribution. For FirstTry, we defined cutpoints at age 22 for the bottom of the distribution and age 30 for the top. A difficulty is involved in defining probands (the low-fertility respondents) for the NumCh variable. Twenty-three percent of the sample had no children. To use these as probands, however, left no variance in the DV in Eq. (1) (because all observations had a value of 0 on the DV). As a result, we used the FirstTry cutpoints to define the probands for the extreme analysis of both variables. High-fertility probands are those who started trying to get pregnant by age 22; low-fertility probands are those who did not start trying to get pregnant until age 30 or later.

The AE (genetic only) model was consistently the best. We present those results in Table 3, which shows a comparison of the (narrow-sense additive) heritabilities from Table 2 for the overall distribution, those from the high-fertility probands, and those from the low-fertility probands. The pattern for the high-fertility heritabilities was fairly similar to that for the overall sample for NumCh. For FirstTry, it was notably lower for males and higher for females. The low-fertility heritability was similar to those from the overall distribution for males, but systematically lower for females, both for FirstTry and especially for NumCh.

In the next analysis, we computed coefficients of variation associated with these fertility variables (see Eq. (8)). Roughly, these coefficients can be interpreted as the part of the estimated phenotypic standard deviation associated with genetic influence, per unit of phenotypic mean (multiplied by 100). Heritabilities, phenotypic mean and standard deviation estimates, and the genetic coefficient of variation CV_a are presented in Table 4. The male and female CVs for NumCh were almost as similar as the male and female heritabilities: the male CV = 39 and the female CV = 35. For FirstTry, CVs once again were very similar for males and

TABLE 3. FITS OF DF ANALYSIS AE (GENETIC-ONLY) MODEL (EQ. (2)) TO DANISH TWIN DATA FOR 35- TO 41-YEAR-OLD TWINS FOR THE WHOLE DISTRIBUTION,[a] HIGH FERTILITY,[b] AND LOW FERTILITY[c]

	Whole Sample		High-Fertility Sample		Low-Fertility Sample	
	h^2	N	h^2	N	h^2	N
Number of Children (NumCh)						
Male-male pairs	.28	1,678	.29	188	.32	380
Female-female pairs	.29	1,540	.26	360	.06	224
Age at First Pregnancy Try (FirstTry)						
Male-male pairs	.35	1,010	.13	161	.23	282
Female-female pairs	.53	1,052	.70	308	.29	173

[a]See Table 1.

[b]FirstTry scores below age 23, from the upper 18.2% of the FirstTry distribution.

[c]FirstTry scores above age 30, from the upper 18.2% of the FirstTry distribution.

for females: male CV = 9, and female CV = 13. In the extreme-fertility analyses for NumCh, male CV = 22 and female CV = 22 for the high-fertility tail; male CV = 34 and female CV = 17 for the low-fertility tail. Genetic CVs were quite low for the FirstTry variable in the upper and lower tails.

SEM Analysis

Using Mx and its maximum-likelihood estimation routine, we refitted the models described above. In the first analysis, we fitted ACE, AE, and ADE models to the twin correlations from Table 2. We report results using single-entry correlations (although both the correlations and the results were almost identical in single-entry and double-entry settings). We used the AIC in selecting the best-fitting model.

Parameter estimates from the ACE, AE, and ADE models were virtually identical to those estimated by DF analysis. Figure 1 shows the basic ACE model that was estimated; we made appropriate adjustments to estimate the AE and ADE models. In Table 5, we compare the SEM results with the results of the DF analysis from the first row of Table 2. According to the AIC criterion, the AE model was the best-fitting model for both male-male and female-female pairs; a maximum-likelihood chi-square showed an excellent match between the model and the data in all cases.

We replicated the DF cross-variable analysis from the bottom line of Table 2, using Mx with the model shown in Figure 2. Parameter estimates for h^2 and c^2 were identical to those from the DF analysis to two decimal places, for both male-male and female-female pairs. SEM provides the flex-

TABLE 4. HERITABILITIES AND COEFFICIENT OF VARIATION FOR GENETIC VARIANCE, 35- TO 41-YEAR-OLD DANISH TWINS

	h^2	M_p	S_p	CV_a
Male-Male, NumCh	.28	1.49	1.10	39
Female-Female, NumCh	.29	1.61	1.06	35
Male-Male, FirstTry	.35	27.22	4.33	9
Female-Female, FirstTry	.53	25.41	4.37	13
Male-Male, High Fertility, NumCh	.29	2.22	.89	22
Female-Female, High Fertility, NumCh	.26	2.11	.92	22
Male-Male, High Fertility, FirstTry	.13	20.86	1.24	2
Female-Female, High Fertility, FirstTry	.70	20.39	1.46	6
Male-Male, Low Fertility, NumCh	.32	1.46	.87	34
Female-Female, Low Fertility, NumCh	.06	1.27	.88	17
Male-Male, Low Fertility, FirstTry	.23	32.45	2.30	3
Female-Female, Low Fertility, FirstTry	.29	32.32	2.02	3

Note: M_p, S_p, and CV_a refer to the mean, standard deviation, and coefficient of variation for the dependent variable. (The p subscript refers to the phenotype, the a to additive genetic variance; see Eq. (8).)

TABLE 5. COMPARISON OF PARAMETER ESTIMATES FROM DF ANALYSIS/SEM ANALYSIS[a]

	ACE Model		AE Model	ADE Model	
	h^2	c^2	h^2	h^2	d^2
Number of Children					
Male-male pairs	.32/.28	–.03/.00	.28/.28[b]	.22/.23	.07/.06
Female-female pairs	.34/.30	–.04/.00	.29/.28[b]	.23/.21	.07/.10
Age at First Pregnancy Try					
Male-male pairs	.31/.30	.03/.04	.35/.35[b]	.41/.35	–.07/.00
Female-female pairs	.46/.46	.05/.05	.53/.52[b]	.62/.52	–.10/.00

Notes: The ACE Model contains components accounting for additive genetic variance, common environmental variance, and nonshared environmental variance/measurement error; the AE model contains additive genetic variance and nonshared environmental variance/measurement error; the ADE model contains additive genetic variance, dominance genetic variance, and nonshared environmental variance/measurement error.

[a]Each analytical approach estimated parameters using the twin correlations and sample sizes presented in Table 1.

[b]Best-fitting models.

ibility to conduct more sophisticated bivariate analyses, however. In Figure 3 we present a model in which we estimate coefficients that indicate shared variance between FirstTry and NumCh, and unique variance in NumCh. To save space, we represent only the model for one member of the twin pair; the other twin's model is identical, and the two are linked with bidirectional paths between the A and the C components like those in the first two figures. This model allows us to account quantitatively for the role of FirstTry in the genetic/environmental partitioning of variability in NumCh.

By the AIC criterion, the ACE model was the best-fitting model for female-female pairs, improving slightly on the AE model (AIC = –20.75 for the ACE model; AIC = –20.01 for the AE model). For male-male pairs, the AE model was preferable (AIC = –20.80 for the ACE model; AIC = –23.41 for the AE model). Standardized parameter estimates are presented in Figure 3 for females. The standardized parameter estimates for males were very similar to those for females.

These estimates may be squared and summed to answer specific questions. For example, the overall variance in NumCh for females in the best-fitting model (the ACE model) can be broken down into genetic variance ($.15^2 + .48^2 = .25$), shared environmental variance ($.21^2 + .00^2 = .04$), and variance associated with nonshared environment/measurement

FIGURE 2. ACE BIVARIATE MODEL, PREDICTING NUMBER OF CHILDREN (NumCh) FROM AGE AT FIRST ATTEMPT TO GET PREGNANT (FirstTry)

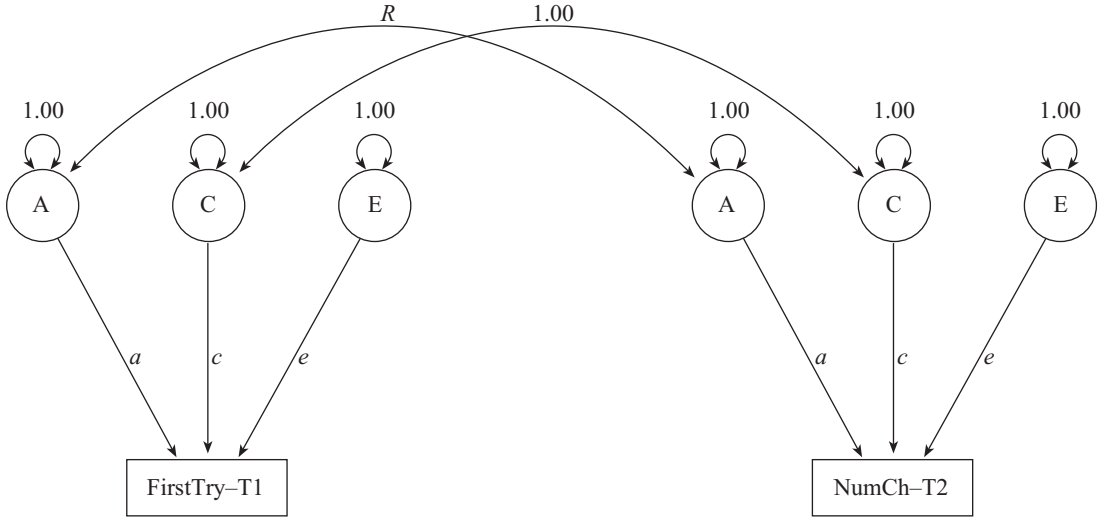
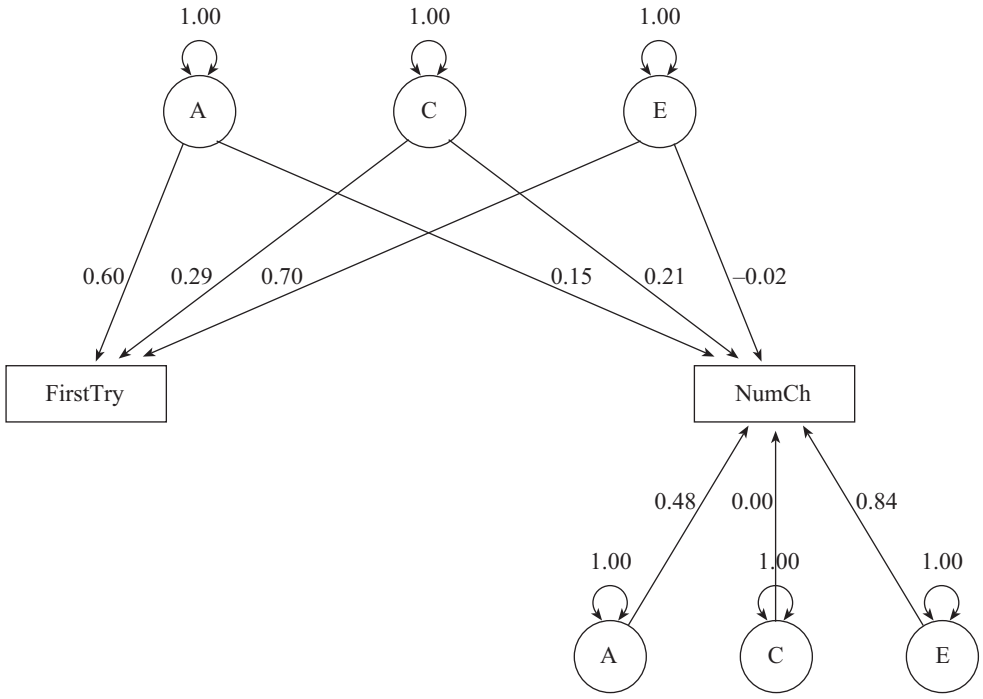


FIGURE 3. ACE BIVARIATE MODEL, SEPARATING VARIANCE IN NUMBER OF CHILDREN (NumCh) THAT OVERLAPS WITH AGE AT FIRST ATTEMPT TO GET PREGNANT (FIRSTTRY) AND THAT IS INDEPENDENT OF FIRST TRY; ESTIMATED PARAMETERS FOR FEMALE-FEMALE PAIRS



error ($-.02^2 + .84^2 = .71$). Of the genetic variance in NumCh, $.48^2 / .25 = .91$ is independent of genetic variance in FirstTry, and $.15^2 / .25 = .09$ is related to genetic variance in FirstTry. For male-male pairs, the best-fitting model (the AE model) defined genetic variance of $.16^2 + .51^2 = .29$ and nonshared environmental variance/measurement error of $-.02^2 + .85^2 = .72$. Of the genetic variance in NumCh, $.51^2 / .29 = .90$ is independent of genetic variance in FirstTry, and $.16^2 / .29 = .09$ is related to genetic variance in FirstTry. To assess the statistical significance of the genetic variance shared between NumCh and FirstTry, we used the difference between chi-squares for the model estimating that term and a model constraining it to be zero. In both cases, removal of this link resulted in a significant decrease in the model's quality of fit (chi-square = 6.11, $df = 1$, $p < .05$ for females; chi-square = 5.25, $df = 1$, $p < .05$ for males).

DISCUSSION

Our analysis of the Danish twin data shows genetic contributions to completed fertility, indicated both by genetic coefficients of variation and by heritabilities. In addition, results across gender categories converge to suggest that slightly more than one-quarter of the variance in completed fertility is attributable to genetic influence.

Further, one of the contributing sources of genetic influence is the age at which individuals first attempt to get pregnant. FirstTry itself contains a major genetic component—higher than that for NumCh, and higher for females than for males. Some of this genetic component is translated to genetic variance in completed fertility.

Both low fertility and high fertility are subject to genetic influence. High fertility has approximately the same level as the overall distribution for NumCh, suggesting that the same genetic influences operate on high fertility as on general fertility for completed fertility. On the other hand, we found lower genetic influence for females on low fertility than on general fertility for both variables, and higher genetic influence for females on high fertility for FirstTry. Both of these findings make substantial theoretical sense, because females waiting until age 30 or older to first try to get pregnant would hardly contribute to reproductive fitness at a level that could maintain genetic variance on this tendency. This perspective would lead us to expect relatively low heritability for low fertility, and high heritability for high fertility for females, exactly the result we obtained. Males, whose fertility behavior is not constrained by age, should have heritabilities that change less sensitively across the high- and low-fertility samples,

especially in the FirstTry variable. These are the patterns that emerged in Table 3.

The analysis of coefficients of variation showed several interpretable patterns. Because heritabilities are proportions, influenced by both phenotypic and genetic variance, they are not directly interpretable across domains. (This theoretical point was made by Fisher 1930 and by numerous others after Fisher; e.g., Houle 1992.) Coefficients of variation are more directly interpretable. If FirstTry is one of several or many contributors to both overall and genetic variance in fertility outcomes, then overall genetic variation in the fertility outcome NumCh should be substantially higher than that in FirstTry. This is demonstrated clearly in that CVs are three to four times as high for NumCh as for FirstTry. FirstTry, however, has CVs of meaningful size, suggesting that variance in first attempt to get pregnant contributes some genetic variance to fertility outcomes. (This finding was supported by our bivariate DF and SEM analyses.)

These results are consistent with Houle's (1992) extensive analysis of fitness traits, in which he found higher genetic variance for traits, the closer they were to directly measuring fitness. Further, he noted that heritabilities and CVs often are negatively related. As he predicted, in our analysis the direct measure of reproductive fitness—NumCh—had lower heritability, but higher CV_a, than our behavioral measure of fertility motivation, FirstTry.

This finding is interpretable, and leads to a theoretical proposal: genetic influences on fertility outcomes appear to operate—at least in part—through volitional fertility motivation and desires. (Quite probably they also operate through biological processes such as fecundity, although we did not directly investigate that question here.) Miller et al. (1999) developed a theoretical rationale for this position, tied to several genetic loci (particularly dopamine receptor genes). If this is true, the proportion of genetic influence on FirstTry would need to be relatively high to account for the fact that its influence will be dampened substantially by other intervening processes. This in fact is exactly the pattern we obtained. If FirstTry is a reliable measure of fertility motivation—a defensible position, especially in a culture like Denmark's, where contraceptive, marriage, and childbearing patterns are such that relatively little undesired fertility occurs—then the results identify a volitional/motivational source for the genetic variance portion of the phenotypic variance in fertility outcomes.

Yet our results in Kohler et al.'s (1999) historical analysis caution against assuming that these genetic influences—and explanations of the influences—are fixed across time. The moderate heritabilities found in the current study occurred in a cohort born during a second demographic transition in Denmark, which Kohler et al. demonstrated was accompanied by increasing and higher heritabilities than during more stable fertility regimes. This finding allows us to blend demographic thinking about the fertility transition with the behavior genetic outcomes reported in this paper to produce a speculative theory.

Kohler et al. (1999) noted that periods of changing fertility norms—such as those that occur at the onset of demographic transition—appear to be those in which genetic influences are allowed to express themselves. On the other hand, if norms are strong and limit individual choice over fertility behavior, genetic influences may be relatively stable. The findings reported here—that a fertility precursor related to fertility motivation, measured as first attempt to get pregnant, has a direct link to the genetic variance in fertility outcomes—provide an explanatory mechanism by which this can occur. In pretransition societies, where childbearing effort begins early and is relatively universal, individual differences in fertility motivation (caused, for example, by genetic tendencies) cannot be demonstrated easily, even if they are present. Similarly, in societies that are considerably posttransition, where childbearing norms dictate the propriety of postponing childbearing for education and other competing activities, individual differences in fertility motivation may not emerge naturally, even if they are present. But in settings in which fertility norms are breaking down, in which individuals have a great deal of latitude and flexibility regarding the number of children they may appropriately have, fertility motivation may play a critical role in defining those outcomes. In such settings, it appears that genetic influences on fertility motivation—and subsequently on fertility outcomes—may emerge and become important.

The findings reported here also match those from a slightly younger cohort in the United States. Rodgers and Doughty (2000) analyzed fertility expectations and (truncated) fertility outcomes among adults age 21 to 28 in kinship data from the U.S. National Longitudinal Survey of Youth. There, fertility outcomes had moderate heritabilities and high CV_as, whereas fertility expectations had higher heritabilities but lower CV_as; this pattern closely matches that described in the current paper. The coherence of these two findings, along with direct genetic theorizing suggesting a source of the influence from Miller et al.'s (1999) research, gives reason to substantially reformulate the basic position deriving from the scant past empirical research on genetic influences on human fertility.

Past research suggested that the answer to the question "Do genes influence human fertility?" is simply "No." This answer, however is much too simple to be correct. Our findings suggest that the answer should be "Sometimes they do, and sometimes they don't." Our theoretical suggestions, offered above, can provide a useful starting point. A complete specification of when they do and when they don't, however, will require careful analysis and a blending of knowledge from a number of different behavioral science disciplines, including demography, behavior genetics, molecular biology, and population biology.

REFERENCES

Adams, J., D.A. Lam, A.I. Hermalin, and P. Smouse. 1990. *Convergent Issues in Genetics and Demography*. New York: Oxford University Press.

Bradburn, N.M., J. Rips, and S.K. Shevell. 1987. "Answering Au-

tobiographical Questions: The Impact of Memory and Inference on Surveys." *Science* 236:159.

Burt, A. 1995. "The Evolution of Fitness." *Evolution* 49:1–8.

Christensen, K., O. Basso, K.O. Kyvik, S. Juul, J. Boldsen, J. Vaupel, and J. Olsen. 1998. "Fecundability of Female Twins." *Epidemiology* 9:189–92.

Danmarks Statistics. 1997. *Vital Statistics, 1995.* Copenhagen: Danmarks Statisticks Trykkeri.

DeFries, J. and D. Fulker. 1985. "Multiple Regression Analysis of Twin Data." *Behavior Genetics* 15:467–73.

Dunne, M.P., N.G. Martin, D.J. Statham, W.S. Slutske, S.H. Dinwiddie, K.K. Bucholz, P.A. Madden, and A.C. Heath. 1997. "Genetic and Environmental Contributions to Variance in Age at First Sexual Intercourse." *Psychological Science* 8:1–6.

Edwards, A.W. 1994. "The Fundamental Theorem of Natural Selection." *Biological Review* 69:443–74.

Ewens, W.J. 1989. "An Interpretation and Proof of the Fundamental Theorem of Natural Selection." *Theoretical Population Biology* 36:167–80.

Falconer, D.S. 1981. *Introduction to Quantitative Genetics*. New York: Longman.

Fisher, R.A. 1918. "The Correlation Between Relatives on the Supposition of Mendelian Inheritance." *Transactions of the Royal Society of Edinburgh* 52:399–433.

———. 1930. *The Genetical Theory of Natural Selection*. Oxford: Clarendon.

Frank, S.B. and M. Slatkin. 1992. "Fisher's Fundamental Theorem of Natural Selection." *Trends in Ecology and Evolution* 7:92–95.

Helweg-Larsen, K., L.B. Knudsen, and B. Petersson. 1998. "Women in Denmark—Why Do They Die So Young? Risk Factors for Premature Death." *Scandinavian Journal of Social Welfare* 7:266–76.

Houle, D. 1992. "Comparing Evolvability and Variability of Quantitative Traits." *Genetics* 130:195–204.

Hughes, K.A. and M.H. Burleson. 2000. "Evolutionary Causes and Consequences of Variation in Fertility and Other Fitness Traits." Pp. 7–33 in *Genetic Influences on Human Fertility and Sexuality,* edited by J.L. Rodgers, D.C. Rowe, and W.B. Miller. Boston: Kluwer.

Imaizumi, Y., M. Nei, and T. Furusho. 1970. "Variability and Heritability of Human Fertility." *Annals of Human Genetics* 33:251–59.

Kohler, H.-P. and J.L. Rodgers. 1999. "DF Analyses of Binary, Ordered, and Censored Variables Using Probit and Tobit Approaches." *Behavior Genetics* 29:221–32.

Kohler, H.-P., J.L. Rodgers, and K. Christensen. 1999. "Is Fertility Behavior in Our Genes? Findings From a Danish Twin Study." *Population and Development Review* 25:253–88.

Kyvik, K.O., A. Green, and H. Beck-Nielsen. 1995. "The New Danish Twin Register: Establishment and Analysis of Twinning Rates." *International Journal of Epidemiology* 24:589–96.

LaBuda, M.C. and J.C. DeFries. 1990. "Genetic Etiology of Reading Disability: Evidence From a Twin Study." Pp. 47–76 in *Perspectives on Dyslexia,* vol. 1, edited by G.T. Pavlidis. New York: Wiley.

LaBuda, M.C., J.C. DeFries, and D.W. Fulker. 1986. "Multiple Regression Analysis of Twin Data Obtained From Selected Samples." *Genetic Epidemiology* 3:425–33.

Lessard, S. 1997. "Fisher's Fundamental Theory of Natural Selection Revisited." *Theoretical Population Biology* 52:119–36.

Lopez, A.D. and R.L. Cliquet. 1984. *Demographic Trends in the European Region: Health and Social Implications*. Albany, NY: World Health Organization.

Lykken, D. and A. Tellegen. 1996. "Happiness Is a Stochastic Phenomenon." *Psychological Science* 7:186–89.

Martin, N.G. and S.A. Treloar. 1991. "Age at Menarche and Fitness: Reply to Reed." *American Journal of Human Genetics* 48:421–23.

Maxwell, S.E. and H.D. Delaney. 1990. *Designing Experiments and Analyzing Data: A Model Comparison Approach*. Belmont, CA: Wadsworth.

Mayo, O.R, R. Burger, and C.R. Leach. 1990. "The Heritability of Fitness: Some Single Gene Models." *Theoretical and Applied Genetics* 79:278–84.

McGue, M. and D.T. Lykken. 1992. "Genetic Influence on Risk of Divorce." *Psychological Science* 3:368–72.

Mealey, L. and N.L. Segal. 1993. "Heritable and Environmental Variables Affect Reproduction-Related Behaviors, But Not Ultimate Reproductive Success." *Personality and Individual Differences* 14:783–94.

Miller, W.B. 1992. "Personality Traits and Developmental Experiences as Antecedents of Childbearing Motivation." *Demography* 29:265–85.

Miller, W.B., D. Pasta, J. MacMurray, W. Chiu, and D. Comings. 1999. "Genetic Influences on Childbearing Motivation and Parental Satisfaction: A Theoretical Framework and Some Empirical Evidence." Pp. 53–102 in *Advances in Population: Psychosocial Perspectives*, vol. 3, edited by L. Severy and W. Miller. London: Jessica Kingsley.

Neale, M.C., S.M. Boker, G. Xie, and H.H. Maes. 1999. *Mx: Statistical Modeling*. 5th ed. Richmond, VA: Medical College of Virginia, Department of Psychiatry.

Neale, M.C. and L.R. Cardon. 1992. *Methodology for Genetic Studies of Twins and Families*. London: Kluwer.

Newcomer, S. 1994. "Research on Adolescent Fertility." Pp. 75–88 in *Advances in Population Psychology: Psychosocial Perspectives,* vol. 2, edited by L. Severy. London: Jessica Kingsley.

Perusse, D., M.C. Neale, A.C. Heath, and L.J. Eaves. 1994. "Human Parental Behavior: Evidence for Genetic Influence and Potential Implications for Gene-Culture Transmission." *Behavior Genetics* 24:327–35.

Plomin, R. 1990. *Nature and Nurture: An Introduction to Human Behavioral Genetics*. Pacific Grove, CA: Brooks/Cole.

Plomin, R., R. Corley, J.C. DeFries, and D.W. Fulker. 1990. "Individual Differences in Television Viewing in Early Childhood: Nature as Well as Nurture." *Psychological Science* 1:371–77.

Plomin, R., J. DeFries, and G.E. McClearn. 1990. *Behavioral Genetics: A Primer*. New York: Freeman.

Plomin, R. and R. Rende. 1991. "Human Behavioral Genetics." Pp. 161–90 in *Annual Review of Psychology,* edited by M.R. Rosenzweig and L.W. Porter. Palo Alto, CA: Annual Reviews Inc.

Price, G.R. 1972. "Fisher's 'Fundamental Theorem' Made Clear." *Annals of Human Genetics* 36:129–40.

Rodgers, J.L. and D. Doughty. 2000. "Genetic and Environmental Influences on Fertility Expectations and Outcomes Using NLSY Kinship Data." Pp. 85–105 in *Genetic Influences on Human Fertility and Sexuality,* edited by J.L. Rodgers, D.C. Rowe, and W.B. Miller. Boston: Kluwer.

Rodgers, J.L., D.F. Harris, and K.B. Vickers. 1992. "Seasonality of First Coitus in the United States." *Social Biology* 39:1–14.

Rodgers, J.L., K. Hughes, H.-P. Kohler, K. Christensen, D. Doughty, D.C. Rowe, and W.B. Miller. Forthcoming. "Genetic Influences Help Explain Variation in Human Fertility Outcomes: Evidence From Recent Behavioral and Molecular Genetic Studies." *Current Directions in Psychological Science.*

Rodgers, J.L. and M. McGue. 1994. "A Simple Algebraic Demonstration of the Validity of the DeFries-Fulker Analysis in Unselected Samples With Multiple Kinship Levels." *Behavior Genetics* 24:259–62.

Rodgers, J.L., D.C. Rowe, and M. Buster. 1999. "Nature, Nurture, and First Sexual Intercourse in the USA: Fitting Behavioral Genetic Models to NLSY Kinship Data." *Journal of Biosocial Sciences* 31:29–41.

Rodgers, J.L., D.C. Rowe, and C. Li. 1994. "Beyond Nature Versus Nurture: DF Analysis of Nonshared Influences on Problem Behaviors." *Developmental Psychology* 30:374–84.

Rossi, A.S. 1994. "Eros and Caritas: A Biopsychosocial Approach to Human Sexuality and Reproduction." Pp. 3–38 In *Sexuality Across the Life Course,* edited by A.S. Rossi. Chicago: University of Chicago Press.

Rowe, D.C. and J.L. Rodgers. 1991. "An 'Epidemic' Model of Adolescent Sexual Intercourse: Applications to National Survey Data." *Journal of Biosocial Science* 23:211–19.

Scarr, S. and S. Grajek. 1982. "Similarities and Differences Among Siblings." Pp. 357–81 in *Sibling Relationships: Their Nature and Significance Across the Lifespan,* edited by M. Lamb and B. Sutton-Smith. Hillsdale, NJ: Erlbaum.

Shryock, H.S. and J.S. Siegel. 1976. *The Methods and Materials of Demography*. New York: Academic Press.

Udry, J.R. 1995. "Sociology and Biology: What Biology Do Sociologists Need to Know?" *Social Forces* 73:999–1010.

Udry, J.R. and B.C. Campbell. 1994. "Getting Started on Sexual Behavior." Pp. 187–208 in *Sexuality Across the Life Course,* edited by A. Rossi. Chicago: University of Chicago Press.

van de Kaa, D.J. 1987. "Europe's Second Demographic Transition." *Population Bulletin* 42:1–57.

Waller, N.B. 1994. "A DeFries and Fulker Regression Model for Genetic Nonadditivity." *Behavior Genetics* 24:149–53.

Waller, N.G., B.A. Kojetin, T.J. Bouchard, D.T. Lykken, and A. Tellegen. 1990. "Genetic Environmental Values: A Study of Twins Reared Apart and Together." *Psychological Science* 1:138–42.

Williams, L.A. and B.J. Williams. 1974. "A Re-examination of the Heritability of Fertility in the British Peerage." *Social Biology* 21:225–31.

Wood, J. 1994. *Dynamics of Human Reproduction: Biology, Biometry, Demography*. Hawthorne, NY: Aldine.