J. biosoc. Sci. (1999) **31**, 29–41

NATURE, NURTURE AND FIRST SEXUAL INTERCOURSE IN THE USA: FITTING BEHAVIOURAL GENETIC MODELS TO NLSY KINSHIP DATA

JOSEPH LEE RODGERS*, DAVID C. ROWE† AND MAURY BUSTER*

**Department of Psychology, University of Oklahoma, Norman, OK 73019, USA and*
†Division of Family Studies, University of Arizona, Tucson, AZ 85721, USA

Summary. Fisher (1930) presented both theoretical and empirical results concerning genetic influences on fertility. Since then, only sparse research has been done on the genetics of fertility, although more sophisticated methodogy and data now exist than were available to Fisher. This paper presents a behavioural genetic analysis of age at first intercourse, accounting for genetic, shared environmental, and selected non-shared environmental influences. The data came from the nationally representative National Longitudinal Survey of Youth (NLSY). A newly developed kinship linking procedure was used that identifies links for cousins, half-siblings, full-siblings and twins in the NLSY. The results suggest a genetic influence in the overall dataset, and also among whites and in male–male and opposite-sex pairs. Genetic influences were extremely small or non-existent for blacks and for female–female pairs. Shared environmental influences were small for most subsets of the data, but moderate for female–female pairs. Two specific non-shared environmental influences – self-esteem and locus of control – were ruled out as accounting for any meaningful variance, although other general sources of non-shared environmental influence appear potentially important. Analysis of selected samples from upper and lower tails suggested that genetic influences are important in accounting for both early and late non-virginity. These findings are consistent with work reported by Miller *et al.* (1999), who used molecular genetic methods. Generally, these findings support the existence of genetic influences and implicate non-shared environmental influences as being important determinants of the timing of loss of virginity among US adolescents and young adults.

Introduction

Hundreds of articles have been published on the influence of social and cultural institutions on adolescent sexual behaviour. Such influences include parents (Newcomer & Udry, 1984), siblings (Rodgers & Rowe, 1988), the community (Billy, Brewster & Grady, 1994; Hogan & Kitagawa, 1985), the church (Reiss, 1967; Mirande,

1968), friends (Billy & Udry, 1985) and potential sexual partners (Rodgers & Rowe, 1993). But virtually no research has been done on the 'initial investment': the influence of genes on adolescent sexual behaviour. Scarr & McCartney (1983) proposed an interactive gene–environment theory suggesting that children and adolescents engage in 'niche-picking' behaviour, in which they seek out the environmental settings that their genetic talents and biological inclinations lead them to select. Research on adolescent sexual behaviour has investigated the niches carefully, but without sufficient attention to the role of genes as both an important explanatory source and as background variation.

Newcomer (1994) stated that 'Partners, peers, parents (maybe even genes) and the community all influence [adolescent sexual] behavior' (p. 85). Udry & Campbell's (1994) literature review found only one small study of genetic influences on adolescent sexual behaviour. Martin, Eaves & Eysenck (1977) used a British twin study (with considerable selection bias) and found that 'cultural differences are less important than individual environmental experiences in determining the age of first sexual intercourse' (p. 91). They also found that 'genetical differences, possibly expressed in part through differences in personality and attitudes, predispose an individual to cross this [first intercourse] threshold at an earlier or a later age' (p. 97), but they also noted that none of their models provided a very effective fit to the data. In their sample of British twins born in the 1950s the mean age at first intercourse for females was 20·7 and for males was 19·3, and many experienced first intercourse within marriage. Obviously, this study has limitations for helping us understand genetic and environmental influences on modern adolescents in the United States.

Virtually no research has been done to address the role that genetic influences play in various aspects of adolescent fertility behaviour, or the trade-off between genetic and environmental influences. Udry's work (e.g. Udry *et al.*, 1985; Udry, Talbert & Morris, 1986) suggested an important role of hormonal influences in both male and female sexual behaviour, and biosocial models of adolescent sexuality are becoming increasingly popular (e.g. Hofferth, 1987; Rodgers & Rowe, 1993; Udry, 1988). But behavioural genetic analysis of adolescent sexual behaviour has apparently never been performed.

Fisher (1930) raised doubts as to whether it would be fruitful to search for genetic influences on fertility behaviour. Plomin, DeFries & McClearn (1990), drawing on Fisher (1930) and Falconer (1981), explained that potential changes in relative fitness across generations due to a particular trait can be measured by the amount of additive genetic variance in that trait present in the population. They concluded that we should 'expect heritability to be low for major components of fitness, such as fertility' (p. 285), and suggested that most genetic variance in such traits should be non-additive.

Nevertheless, Fisher himself was willing to search for genetic variance that his own theoretical analysis showed should not exist. In the same book (Fisher, 1930), in which he questioned whether genes could ever influence fertility, Fisher used correlations of completed family size across three generations of British families from the late 1800s to estimate a significant heritability of $h^2 = 0{\cdot}40$. According to Williams & Williams (1974), this estimate has been 'frequently cited and has served as a stimulus for other studies of the inheritance of fertility' (p. 225). Williams and Williams criticized the estimate, however, as containing spurious correlation caused by secular changes in the date of marriage.

Fisher's theoretical arguments apply most directly to measures of fertility and family size, the ultimate dependent variables in any discussion of reproductive success. They also apply to any other variable – like age at first intercourse – that has a relationship to fertility outcomes. However, this expectation depends on a long enough period of time that traits with selective advantage can realize that advantage. In a population with natural fertility (i.e. little reliance on effective contraception), age at first intercourse should have a strong relationship to ultimate reproductive success.

This investigation looks at age at first sexual intercourse in the US population using a sample who were adolescents during the 1970s and early 1980s. During the past several centuries, there have been secular changes – both up and down – in age at first intercourse. Furthermore, the availability of reliable contraception must weaken the selective advantage of early onset of sexual behaviour that occurs in societies with little or no contraceptive use. Such changes could certainly act to weaken the selective value of early onset of sexual activity. Given these changes, it is an important theoretical question to ask whether genes currently play a role in influencing onset of sexual behaviour. At the same time, the role of environmental influences is also of particular interest and importance. The modelling in this study simultaneously addresses the role of both types of influence.

The data used come from the National Longitudinal Survey of Youth (NLSY). To separate genetic and environmental influences requires data from different kinship levels, which are not defined explicitly in the NLSY. However, because the household structure of the NLSY contains many unidentified kinship links, a linking algorithm (J. L. Rodgers, unpublished document) was developed that uses NLSY variables to classify kinship pairs into half-sibling, full-sibling, twin and cousin pairs. This kinship structure is used, along with a recently developed regression procedure – DF analysis (DeFries & Fulker, 1985; Rodgers, Rowe & Li, 1994) – to analyse variance in age at first intercourse into that attributable to genetic, shared environmental, and non-shared environmental influences. Because patterns of sexual debut differ substantially across race and genders, the models are fitted separately by these demographic categories. There is a brief description of the NLSY data and the linking algorithm, followed by a short exposition on DF analysis. Then analytic models are fitted to the NLSY kinship data, and results of that fitting routine presented. The paper concludes with a discussion of the implications of these findings.

The NLSY data and the linking algorithm

The data

The NLSY began in 1979 as a household probability sample of all 14- to 21-year-old youths in selected households, with $N = 11{,}406$ civilian respondents. Because of this design, many kinship links exist in the NLSY files, including twins, full-siblings, half-siblings, cousins and adoptive siblings. The original sample has been followed on a yearly basis since 1979, with over 90% continuation through 1992.

In 1983, 1984 or 1985, the NLSY youth (who were 20–27 years old in 1985) retrospectively reported their age at first intercourse. Of the female respondents, 4800 gave reports in both 1983 and 1984, which provided a test–retest reliability estimate for the measure of age at first intercourse. The correlation between these two reports was

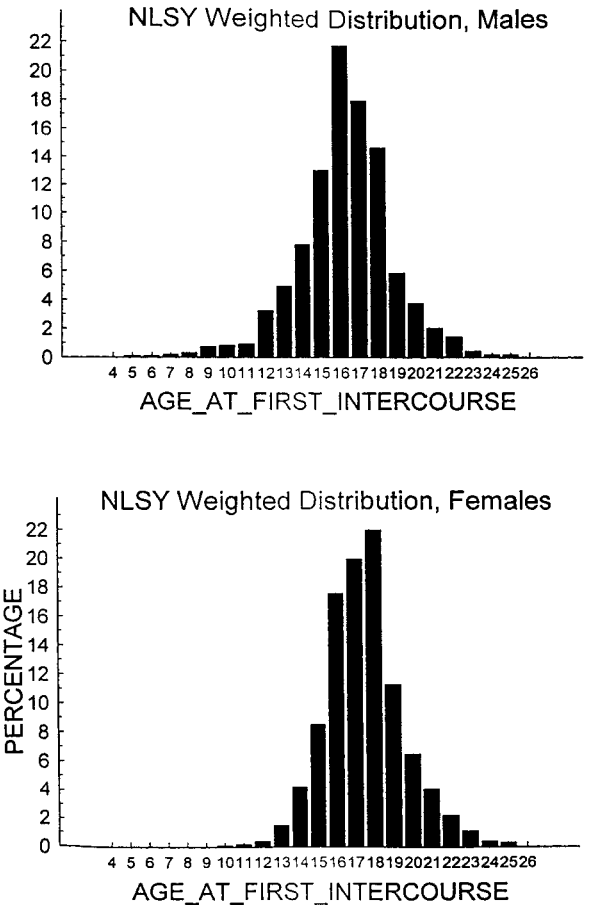

Fig. 1. Distribution of age at first intercourse, estimated from overall NLSY dataset.

$r = 0{\cdot}84$. Eighty-eight per cent of these female respondents gave ages that were exactly the same or within a year, and 96% gave reports that were within 2 years. These results support the reliability of this dependent measure (although male reports are often more problematic than those for females, e.g. Rodgers, 1996). Other reliability and validity analyses of survey data from sexuality studies have also found acceptable psychometric properties of these types of data (e.g. Catania *et al.*, 1990 or Rodgers, Billy & Udry, 1982).

Non-missing reports were averaged across these 3 years when they were not consistent to construct the basic dependent variable. Figure 1 shows the distribution of age at first intercourse from the whole NLSY sample, with weights used to adjust for attrition and oversampling of some subgroups. This distribution estimates frequencies in each category for the whole US population from this time period. Figure 1 shows that there are a few reports of pre-pubertal first intercourse. The distributions were adjusted by dropping all responses of age 10 or less, for both theoretical and empirical reasons. Rodgers (1996) presented empirical analyses showing logical inconsistencies in

https://doi.org/10.1017/S0021932099000292 Published online by Cambridge University Press

Table 1. Means and standard deviations (SD) for overall distribution of age at first intercourse (years) for the NLSY kinship sample, by gender and race

	Mean	SD
Overall sample	16·7	2·6
Males	15·9	2·7
Females	17·7	2·2
Whites	17·2	2·5
Blacks	15·7	2·7
White males	16·6	2·7
White females	18·0	2·2
Black males	14·4	2·6
Black females	17·0	2·1

such reports that raise doubts about their validity. Further, even if these reports are correct, pre-pubertal sexual behaviour has no potential for influencing fertility, which substantially reduces interest in these reports. Virtually all of the reports of first intercourse before the age of 10 came from males. When the frequencies are adjusted by sampling weights, 0·2% of the females and 2·2% of the males reported first intercourse before the age of 10. Table 1 provides first intercourse means (before adjusting for pre-pubertal intercourse) for the sample and subsamples that are used in this study. Independent variables included the genetic coefficient (described in the next section), and measures of self-esteem (collected in 1980) and locus of control (collected in 1979), the latter two of which were used to account for potential non-shared sources of influence due to different personality types.

The linking algorithm

The NLSY sample contains many kinship links because of the household sampling design of the survey. However, no information is available within the NLSY data to distinguish explicitly several ambiguous kinship categories. For example, the 'sibling' category contains a mixture of full-, half- and step/adoptive siblings, whose coefficients of genetic relationship are 0·50, 0·25 and 0·00, respectively. Further, no information is available to distinguish explicitly MZ and DZ twin pairs in the 35 twin pairs in the NLSY sample.

However, a number of specific variables contained in the NLSY files give information to help resolve the ambiguity in these kinship links. The information in these variables can be used, along with a decision algorithm that objectively applies a set of rules for classification, to classify the NLSY respondent pairs into kinship categories. A similar problem exists in the NLSY children data, a sample of all of the children ever born to the NLSY females. Rodgers *et al.* (1994) presented results of a linking algorithm for that data source that was used in a study of the environmental and genetic variance associated with childhood problem behaviours.

https://doi.org/10.1017/S0021932099000292 Published online by Cambridge University Press

Rodgers (unpublished document) gives a careful account of the logic underlying the NLSY linking algorithm. A copy of the documentation, along with a diskette containing the resulting kinship links, can be obtained from the first author of this paper. Only a brief summary of the algorithm is presented here. The algorithm identified 3890 kinship pairs that were targeted for classification. In general, the R coefficient – the coefficient of genetic relatedness for each pair – was defined: $R = 1{\cdot}0$ for MZ twins, $R = 0{\cdot}50$ for DZ twins and full-siblings, $R = 0{\cdot}25$ for half-siblings, $R = 0{\cdot}125$ for cousins, and $R = 0{\cdot}00$ for adoptive siblings. These values give the percentage of shared genes on average in the given kinship category.

Cousins were unambiguously identified in the NLSY files as such. If both members of a kinship pair agreed that they were cousins, the respondent's classification was simply accepted as correct and assigned an $R = 0{\cdot}125$. Twins were assigned on the basis of shared birthdays; however, MZ and DZ twins were not distinguished in the data. All opposite-sex twins were assigned $R = 0{\cdot}5$. Of the remaining same-sex twins, half will be MZ and half DZ in the population. Thus, in the absence of further knowledge, all same-sex twin pairs were classified as $R = 0{\cdot}75$.

The sibling pairs were the most critical because of the large sample size. In 1988 (when they were 23–30 years old), respondents created a retrospective time line from age 0 to age 18 of whether or not they lived with their biological father and their biological mother in each year. This set of variables, along with several other 'living together' variables, were combined into an algorithm to identify kinship level. If there was ambiguity left about level of relatedness, an intermediate kinship level was assigned (e.g. if two related adolescents were clearly either full- or half-siblings, but could not be resolved, the algorithm assigned $R = 0{\cdot}375$). This linking procedure assigned an R coefficient to 2338 of the kinship pairs, a 60% classification rate. By category, these included 76 cousin pairs ($R = 0{\cdot}125$), 43 half-sibling pairs ($R = 0{\cdot}25$), 310 pairs that were ambiguous half-sibling/full-sibling pairs ($R = 0{\cdot}375$), 1877 full-sibling pairs ($R = 0{\cdot}50$), and 32 same-sex twin pairs of unknown zygocity ($R = 0{\cdot}75$).

An additional method was used to increase the sample size, which was still somewhat small for subgroup analyses. In this approach, the height variable was used as a diagnostic variable. The logic underlying this method is that for a highly heritable variable like height, those with high levels of genetic relatedness will in general be more similar than those with lower levels. This logic was implemented with three different values of R coefficient that emerged from the algorithm already described above: $R = 0{\cdot}75$, $R = 0{\cdot}375$ and $R =$ missing. This new dataset was much larger than the more conservative set of links described for the first dataset, particularly because of the addition of many of the links from the R-missing category; it included identified links for 3419 out of the original 3890 pairs, an 88% classification rate. These broke down into 20 MZ twin pairs, 3090 full-sibling pairs, 233 half-sibling pairs, and 76 cousin/adoptive sibling pairs. Primary results presented in this paper came from use of the smaller dataset.

DF analysis, adapted for non-shared influences

DF analysis is a multiple regression technique originally proposed by DeFries & Fulker (1985) that uses kinship pairs with multiple levels of relatedness (e.g. MZ and DZ twins, full-siblings, half-siblings, cousins) to estimate genetic and shared environmental

influences (h^2 and c^2) on a specified trait. The procedure uses measures of a trait from kinship pairs representing at least two levels of genetic relatedness. The following least squares regression equation is estimated:

$$K_1 = b_0 + b_1 K_2 + b_2 R + b_3 (K_2 * R) + e, \tag{1}$$

where K_1 is the score for the first member of the kinship pair, K_2 is the score for the second member of the kinship pair, R is the coefficient of genetic relatedness, the b values are least squares regression coefficients, and e is the error or residual. Under the basic assumptions of the additive genetic model, it can be shown that b_1 gives an unbiased estimate of shared environmental influence (c^2) and b_3 gives an estimate of heritability (h^2) (LaBuda *et al.*, 1986; Rodgers & McGue, 1994). Critical assumptions of the model include additivity, no assortative mating, and equal shared environmental influences across level of genetic relatedness. If this latter assumption is not met, then the c^2 value estimates the average of the shared environmental influences across the genetic categories. The ordering defining which member of the kinship pair is K_1 and which is K_2 is arbitrary; this problem is solved by double-entering the data, so that each individual defines both a K_1 and K_2 score (e.g. LaBuda & DeFries, 1990).

The basic DF analysis model does not account for non-shared environmental influences. If non-shared environmental influences are present, these will be absorbed in the residuals of equation (1). Rodgers *et al.* (1994) showed how to enter signed kinship difference scores to test for specific non-shared environmental influences. Further, the interaction of the difference variable with the genetic coefficient R can be used to assess whether there are significant non-shared genetic influences on the trait as well. The models that are fitted are the following:

$$K_1 = b_4 + b_5 K_2 + b_6 R + b_7 (K_2 R) + b_8 \mathrm{ENVDIF} + e \quad \text{and} \tag{2}$$

$$K_1 = b_9 + b_{10} K_2 + b_{11} R + b_{12} (K_2 R) + b_{13} \mathrm{ENVDIF} + b_{14} (\mathrm{ENVDIF} * R) + e, \tag{3}$$

where ENVDIF is a signed difference score from the two kin on a specific measured environmental source that might account for differences between them on the trait, and the variables, parameters and residual are defined as before. These types of processes can be explicitly accounted for while estimating h^2 and c^2 simultaneously within the same model. In this paper this type of model was fitted to age at first intercourse scores from related individuals in the NLSY dataset to estimate the level of heritability and shared environment that underlie these scores. In addition, measures of personality differences were entered – self-esteem and locus of control – into equations (2) and (3) above. If these difference variables were found to be significant, then differences in these personality traits can be considered to underlie differences in delinquency level, after accounting for genetic and environmental similarity.

One additional DF analysis was run to account for heritability and shared environmental variance related to extreme values of the first intercourse variable. DeFries & Fulker's (1985) first DF analysis was applied to a selected sample in which respondents were selected for a reading disability. The genes and shared environmental influences that cause someone to be extreme on a trait may be very different from those accounting for individual differences within a whole distribution (LaBuda & DeFries, 1990). For example, a single gene may cause a reading disability, while general reading

Table 2. Heritabilities (h^2) and shared environmental variance (c^2) (with standard errors indicated) for the NLSY kinship data, by race and gender subgroups

	h^2	c^2
Whites ($N=1322$)	$0{\cdot}51 \pm 0{\cdot}39$*	$-0{\cdot}02 \pm 0{\cdot}18$
Blacks ($N=469$)	$0{\cdot}09 \pm 0{\cdot}47$	$0{\cdot}15 \pm 0{\cdot}23$
Male–male ($N=544$)	$0{\cdot}54 \pm 0{\cdot}48$	$0{\cdot}09 \pm 0{\cdot}24$
Female–female ($N=497$)	$0{\cdot}15 \pm 0{\cdot}51$	$0{\cdot}27 \pm 0{\cdot}25$
Opposite sex ($N=866$)	$0{\cdot}38 \pm 0{\cdot}47$	$-0{\cdot}02 \pm 0{\cdot}23$
Overall dataset ($N=1907$)	$0{\cdot}37 \pm 0{\cdot}29$*	$0{\cdot}08 \pm 0{\cdot}14$

*Significant, $\alpha = 0{\cdot}10$, sample sizes adjusted back to number of pairs; see text for rationale. N = number of kinship pairs.

ability may be a highly polygenic trait. This same reasoning applies to first intercourse as well; the genes or shared environmental influences that lead to extremely early or extremely late first intercourse may be different from those defining general individual differences. Thus, analyses were run on the bottom and top 15–20% of the distributions, to account for the possibility of different patterns in these selected samples.

Results

Table 2 shows heritability (h^2) and shared environmental variance (c^2) for the NLSY kinship data. Heritabilities were moderate to large for the overall dataset and for whites, and small for blacks. The c^2 values were consistently small, suggesting little or no shared environmental influences. The analysis by gender categories suggested a fairly substantial heritability for male–male pairs, small for female–female pairs and moderate for opposite-sex pairs. Shared environmental variance was small for male–male and opposite-sex pairs and moderate for female–female pairs. When the analysis was run for opposite-sex pairs, the mean gender difference was adjusted out (see Table 1) by subtracting the difference between the male and female means from each female's score.

The DF analysis model provides point estimates of h^2 and c^2. An appropriate statistical test is difficult to define, both because of the double entry (which requires adjustment back to a smaller sample size to account for independent of errors; the usual approach of using the original sample size – half of the double-entered sample – is a conservative method for making this adjustment) and because of the low power associated with a multiplicative interaction (e.g. Aiken & West, 1991) like the variable (K_2*R) interaction term used to estimate h^2 (see equation 1). Standard errors and statistical tests were adjusted back to the original sample size of pairs, but a less stringent $\alpha = 0{\cdot}10$ was adopted in evaluating h^2 and c^2. Two statistically significant findings with this criterion were for h^2 in the overall dataset and the white sample; a number of other results were, in addition, suggestive of genetic influence.

https://doi.org/10.1017/S0021932099000292 Published online by Cambridge University Press

Table 3. Means, standard deviations (SD), heritabilities (h^2), and shared environmental variances (c^2) for the upper and lower tails for whites, blacks and total from the NLSY age at first intercourse distributions

	Bottom	Top
Whites ($N = 860$ in lower tail, 724 in upper tail)		
Mean	14·1	20·0
SD	1·0	1·7
h^2	0·34	0·63
c^2	−0·14	−0·25
Blacks ($N = 457$ in lower tail, 404 in upper tail)		
Mean	13·0	18·4
SD	1·0	1·5
h^2	0·34	−0·04
c^2	−0·18	0·02
Total dataset ($N = 1043$ in lower tail, 978 in upper tail)		
Mean	13·2	19·9
SD	1·3	1·9
h^2	0·15	0·43
c^2	−0·06	−0·16

All analyses presented in Table 2 were re-run in three different ways, to evaluate the stability of these findings under different reasonable specifications. First the larger dataset that classified 88% of the total kinship pairs was used. There were a few anomalies (e.g. a negative heritability for blacks), but results supported the existence of genetic influences for the total dataset, whites and male–male pairs. Second, the original dataset was used, deleting the cousin pairs. Cousin pairs were deleted both because they are more likely to violate the equal-environments assumption of the additive genetic model, and also because the kinship correlations for the cousins were larger than for other genetic categories, suggesting the possibility of misclassifications in this category. For whites, heritabilities were slightly higher for these categories than in the original dataset ($h^2 = 0{\cdot}74$ for whites, $h^2 = 0{\cdot}11$ for blacks, and $h^2 = 0{\cdot}38$ for the overall dataset), while c^2 values were consistently around zero in each case. Finally, these analyses were re-run using the original dataset without deleting respondents who reported pre-pubertal first intercourse, and heritabilities slightly lower than those in Table 2 were found ($h^2 = 0{\cdot}29$ for the overall dataset, $h^2 = 0{\cdot}38$ for whites, and $h^2 = 0{\cdot}00$ for blacks). These patterns support the stability of the original findings from Table 2 across different ways of structuring the dataset.

In the next analysis, self-esteem and locus of control were entered into the models, as described in equations (2) and (3). There were no significant results found for these particular non-shared influences, nor for their interactions with the genetic coefficients. This result does not rule out non-shared influences as a general explanatory source, only these specific ones.

https://doi.org/10.1017/S0021932099000292 Published online by Cambridge University Press

Results from the selection analysis are presented in Table 3. To obtain these, the extreme 15–20% at the top and bottom of the overall double-entered distribution were defined for each of the categories, and a (single-entered) DF analysis was run on these extreme samples (some kinship pairs were double-entered, if they happened to both have extreme scores). Because there were many tied scores, a specific percentage criterion for each distribution could not be defined. Rather, all scores were used beyond the extreme that came as close to cutting off 15–20% as possible. For the overall dataset, using first intercourse scores for those with ages less than 15 years cut off the bottom 15·0% of the distribution, and those greater than 19 years cut off the top 14·1%. For whites, ages $<15{\cdot}5$ years cut off 19·2%, and ages >19 years cut off 16·3%. For blacks, ages $<14{\cdot}5$ years cut off 23·3%, and ages greater than 17·5 years cut off the upper 20·6%.

Generally, heritabilities in both the upper and lower tails were of moderate size. Even for blacks, who had very small heritabilities in the unselected analysis, lower tailed h^2 was of moderate size (though essentially zero for the upper tail). The value of c^2 was consistently small or slightly negative. These patterns suggest that both precocious and delayed first intercourse may have genetic bases. The variance still unexplained is attributable to measurement error and unspecified non-shared environmental influences.

Discussion

As discussed in the Introduction, there is little empirical literature on genetic influences on age at first intercourse to guide expectations. A large literature on environmental influences exists (e.g. Hofferth, 1987), but outside the context of the additive genetic model employed in this study. It would be useful to classify such research into that accounting for shared and non-shared environmental influences.

However, there are theoretical reasons to expect age at first intercourse *not* to have genetic aetiology. To the extent that selective processes have acted on this variable, no additive genetic variance of the type that is identified by the DF analysis model would be expected. However, contraceptive efficacy should act to reduce (although not to eliminate) the selective advantage of early sexual debut, and social and cultural changes could also act to create such advantages. For example, if some social innovation suddenly changes what the population views as sexually attractive, this change could cause heritabilities and changing environmental influences until selective processes achieve an equilibrium again. It is in this context that a behavioural genetic study of age at first intercourse is of interest.

Are there genetic influences acting on age at first intercourse? The results of this study suggest that there are, but with certain qualifications. Blacks, for example, appear to have little genetic or shared environmental variance to account for. Female–female pairs were more influenced by the shared environment than by genetic sources. But consistent and meaningful heritabilities were found, particularly among whites and in the overall distribution, and also for male–male pairs and for opposite-sex pairs. The shared environmental influences were never large or consistent enough to be strongly interpretable. This may be an important finding, given the particular fascination the social science literature has had with developing socialization models to account for sexuality variables like age at first intercourse. The largest c^2 found was for female–female pairs, which matches results from the literature. For example, Billy

and his colleagues found influence models to be applicable to female friendship pairs, but not male friendship pairs (e.g. Billy, Rodgers & Udry, 1984; Billy & Udry, 1985).

Further, the finding that genetic influences appear to operate in the extremes of the first intercourse distributions at similar levels to the overall distribution suggests that the genes contributing to early and late first intercourse operate very similarly to those defining individual differences across the whole distribution (and may well be the same genes). This finding has important practical implications for the delivery of programmes – sex education, pregnancy prevention, etc. – to adolescents and young adults of different ages.

The finding that a great deal of variance is left over to attribute to the non-shared environment has a number of important implications. First, it leaves intact and uncontaminated the large body of research investigating environmental influences on early adolescent sexual behaviour (although some of that research has clearly sought *shared* environmental sources of explanation). Such non-shared influences include differential parental treatment and different friendship networks. Not even twins will share exactly the same friendship circles, yet it is exactly that circle that provides the potential sexual partners. Two non-shared sources that have been explicitly identified as important in past research include differential maturation rates across birth order (Rodgers, Rowe & Harris, 1992) and social contagion processes in which potential partners influence virgin adolescents to have sexual intercourse (Rodgers & Rowe, 1993). A second implication of the importance of non-shared environmental influences is to support the value of intervention programmes (e.g. school sex education and clinic-based health programmes).

The results of this study are among the first to identify a clear role for genetic influences in accounting for individual differences in age at first intercourse. However, it is not yet clear why black pairs had low genetic aetiology compared with other categories, although the absence of genetic effect doesn't imply that selective processes have disappeared (Fisher's work in fact shows the reverse to be true). This discussion has already speculated about why there would be significant heritabilities in the several categories in which they were identified. Certainly, contraceptive development is the most obvious explanation for why this important fertility variable should violate Fisher's theorem. Other possible explanations – that are not independent of contraceptive practices – include increased societal efforts to exert control over sexual behaviour (e.g. 'Just Say No' programmes), the availability of abortion, sex education programmes in schools, and secular changes in social contagion processes (Rodgers & Rowe, 1993). To the extent that these sorts of innovations affect age at first intercourse in either direction, they can also affect the balance between h^2, c^2 and non-shared environmental variance, which the DF model estimates. It is an important theoretical point that heritability does not exist in a vacuum, but rather accounts for variance in relationship to other theoretical sources from the environment (although Fisher's theorem applies to the presence of additive genetic variance in a broad sense).

As an important coincidence, when the current paper was presented at a conference, a paper by Miller *et al.* (1999) was also presented that used the same dependent variable – age at first intercourse – but a completely different methodology to account for genetic influences. Miller and colleagues used molecular genetic techniques to demonstrate a relatioship between dopamine receptors and age at first intercourse,

using a sample of 414 adult white men and women. The dopaminergic system has been previously identified as playing a critical role in regulating sexual behaviour in humans. The conclusion from each of these two papers – that genes can account for important variance in age at first intercourse – obtained from two completely different methodological approaches and two different types of samples, substantially strengthens the individual conclusions from these two studies. The present study's strongest finding on race – that genetic influences operate especially among the white sample – perfectly matches Miller *et al.*'s results, which were found for an entirely white sample. Miller *et al.*'s paper speculates that there may be race differences in dopamine-related genotypes, which could help to elucidate the clear differences found between blacks and all of the other racial categories in the current paper. Further, Miller *et al.* also found a stronger genetic basis to age at first intercourse for males than for females, a finding that exactly matches the pattern in Table 2 of this paper. Such findings in independent studies using different methodologies certainly imply that the findings are meaningful and generalizable.

In conclusion, there is a need for more research. Other fertility variables – frequency of intercourse, abortion rates and contraceptive use, among others – should also be investigated to elucidate further the effects of genetic and environmental variables on fertility behaviours. Despite the validity of Fisher's theorem in idealized settings like those defined by his assumptions, there is still considerable interest in the questions that he put to rest theoretically in more ecologically valid settings. Further work on the genetic and environmental influences on fertility behaviours is, of course, certainly needed.

Acknowledgments

This research was supported by Grant #R01-H02-1973 from the National Institutes of Health.

References

Aitken, L. S. & West, S. G. (1991) *Multiple Regression: Testing and Interpreting Interactions.* Sage, Newbury Park, CA.

Billy, J. O. G., Brewster, K. L. & Grady, W. R. (1994) Contextual effects on the sexual behavior of adolescent women. *J. Marriage Fam.* **56,** 387–404.

Billy, J. O. G., Rodgers, J. L. & Udry, J. R. (1984) Adolescents' sexual behavior and friendship choice. *Social Forces* **62,** 653–678.

Billy, J. O. G. & Udry, J. R. (1985) Patterns of adolescent friendship and effects on sexual behavior. *Social Psychol. Q.* **48,** 27–41.

Catania, J. A., Gibson, D. R., Chitwood, D. D. & Coates, T. J. (1990) Methodological problems in AIDS behavioral research: Influence of measurement error and participation bias in studies of sexual behavior. *Psychol. Bull.* **108,** 339–362.

DeFries, J. C. & Fulker, D. W. (1985) Multiple regression analysis of twin data. *Behav. Genet.* **15,** 467–473.

Falconer, D. (1981) *Introduction to Quantitative Genetics.* Longman, London.

Fisher, R. A. (1930) *The Genetical Theory of Natural Selection.* Dover Publishers, New York.

Hofferth, S. L. (1987) Factors affecting initiation of sexual intercourse. In: *Risking the Future: Adolescent Sexuality, Pregnancy, and Childbearing – Working Papers and Appendixes.* Edited by S. L. Hofferth & C. D. Hayes. National Academy Press, Washington.

Hogan, D. P. & Kitagawa, E. M. (1985) The impact of social status, family structure, and neighborhood on the fertility of black adolescents. *Am. J. Sociol.* **90,** 825–855.

Martin, N. G., Eaves, L. J. & Eysenck, H. J. (1977) Genetical, environmental, and personality factors influencing the age of first sexual intercourse in twins. *J. biosoc. Sci.* **9,** 91–97.

Miller, W. B., Pasta, D. J., MacMurray, J., Chiu, C., Wu, H. & Comings, D. E. (1999) Dopamine receptors are associated with age at first sexual intercourse. *J. biosoc. Sci.* **31,** 43–54.

Mirande, A. M. (1968) Reference group theory and adolescent sexual behavior. *J. Marriage Fam.* **30,** 572–577.

Newcomer, S. (1994) Research on adolescent fertility. In: *Advances in Population Psychology: Psychosocial Perspectives.* Edited by L. Severy. Jessica Kingsley.

Newcomer, S. & Udry, J. R. (1984) Mothers' influence on the sexual behavior of their teenage children. *J. Marriage Fam.* **46,** 477–485.

LaBuda, M. C. & DeFries, J. C. (1990) Genetic etiology of reading disability: Evidence from a twin study. In: *Perspectives on Dyslexia*, Vol. 1, pp. 47–76. Edited by G. T. Pavlidis. Wiley, New York.

LaBuda, M. C., DeFries, J. C. & Fulker, D. W. (1986) Multiple regression analysis of twin data obtained from selected samples. *Genet. Epidemiol.* **3,** 425–433.

Plomin, R., DeFries, J. C. & McClearn, G. E. (1990) *Behavioral Genetics: A Primer.* Freeman, New York.

Reiss, I. L. (1967) *The Social Context of Premarital Sexual Permissiveness.* Holt, Rinehart and Winston, New York.

Rodgers, J. L. (1996) Sexual transitions in adolescence. In: *Transitions through Adolescence. Interpersonal Domains and Context.* Edited by J. A. Graber, J. Brooks-Gunn & A. C. Peterson. Erlbaum, Mahwah, NJ.

Rodgers, J. L., Billy, J. O. G. & Udry, J. R. (1982) The recission of behaviors: Inconsistent responses in adolescent sexuality data. *Social Sci. Res.* **11,** 280–296.

Rodgers, J. L. & McGue, M. (1994) A simple algebraic demonstration of the validity of DeFries–Fulker analysis in unselected samples with multiple kinship levels. *Behav. Genet.* **24,** 259–262.

Rodgers, J. L. & Rowe, D. C. (1988) Influence of siblings on adolescent sexual behavior. *Devl Psychol.* **24,** 722–728.

Rodgers, J. L. & Rowe, D. C. (1993) Social contagion and adolescent sexual behavior: A developmental EMOSA model. *Psychol. Rev.* **100,** 479–510.

Rodgers, J. L., Rowe, D. C. & Harris, D. F. (1992) Sibling differences in adolescent sexual behaviour: Inferring process models from family composition patterns. *J. Marriage Fam.* **54,** 142–152.

Rodgers, J. L., Rowe, D. C. & Li, C. (1994) Beyond nature vs. nurture: DF analysis of nonshared influences on problem behaviors. *Devl Psychol.* **30,** 374–384.

Scarr, S. & McCartney, K. (1983) How people make their own environments: A theory of genotype–environment effects. *Child Dev.* **54,** 424–435.

Udry, J. R. (1988) Biological predispositions and social control in adolescent sexual behavior. *Am. sociol. Rev.* **53,** 709–722.

Udry, J. R., Billy, J. O. G., Morris, N., Groff, T. & Raj, M. (1985) Serum androgenic hormones motivate sexual behavior in adolescent boys. *Fert. Steril.* **43,** 90–94.

Udry, J. R. & Campbell, B. C. (1994) Getting started on sexual behavior. In: *Sexuality across the Life Course.* Edited by A. Rossi. University of Chicago Press, Chicago.

Udry, J. R., Talbert, L. M. & Morris, N. M. (1986) Biosocial foundations for adolescent female sexuality. *Demography* **23,** 217–230.

Williams, L. A. & Williams, B. J. (1974) A re-examination of the heritability of fertility in the British peerage. *Social Biol.* **21,** 225–231.