

Genetics of Educational Attainment and the Persistence of Privilege at the Turn of the 21st Century*

François Nielsen and J. Micah Roos

June 2011

Abstract

We use structural equations methodology with data on 1,576 pairs of variously-related young adults (MZ twins, 200 pairs; DZ twins, 324 pairs; full siblings, 639 pairs; half siblings, 213 pairs; cousins, 68 pairs; and nonrelated siblings, 132 pairs) to distinguish the roles of genetic and environmental influences on educational attainment (highest degree earned). While many cognitive and educational outcomes show increasing effects of genes (heritability) and declining effects of the shared environment by late adolescence, we hypothesize that shared environment effects may persist into early adulthood for educational attainment, as tertiary education involves financial expenses more directly affected by available family resources than more purely cognitive outcomes. We estimate quantitative genetic (ACE) models of attainment controlling for sex and age. We find that the role of genes in educational attainment is relatively weaker (23 percent of the variance in attainment) and the role of the shared family environment stronger (41 percent of the variance for twins and 30 percent of the variance for nontwin siblings) than is typically found for cognitive outcomes (such as IQ) at the same young adult stage in the life course. The pattern of persistent shared environmentality, especially for twins, is not accounted for by the strong degree of assortative mating in the data (parental correlation $\hat{r} = .629$) nor by direct effects of educational attainment of the siblings on each other. We conclude that this empirical pattern indicates the persistence of substantial inequality of opportunity for educational attainment in American society at the turn of the twenty first century.

1 Introduction

A principal finding of much research on social stratification and mobility is that education is a central mechanism of individual socioeconomic achievement in modern industrial societies. Education plays a dual role. On one hand it is the principal mechanism of social reproduction, producing an association between status of origin and status of destination. On the other it is the principal channel of

*This research uses data from Add Health, a program project directed by Kathleen Mullan Harris and designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris at the University of North Carolina at Chapel Hill, and funded by grant P01-HD31921 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development, with cooperative funding from 23 other federal agencies and foundations. Special acknowledgment is due Ronald R. Rindfuss and Barbara Entwisle for assistance in the original design. Information on how to obtain the Add Health data files is available on the Add Health website (<http://www.cpc.unc.edu/addhealth>). No direct support was received from grant P01-HD31921 for this analysis. We are grateful to Matt Bradshaw, Shawn Bauldry and participants in the Social Stratification Seminar at University of North Carolina at Chapel Hill for useful comments on previous drafts.

social mobility by which individuals may reach status destinations different from their status of origin (Blau and Duncan 1967; Boudon 1974; Hout and DiPrete 2006). However the nature of the processes by which social origins and individual characteristics determine the educational achievement of individuals has remained mysterious, receiving relatively little attention in recent research on social mobility (Hout and DiPrete 2006).

The bulk of sociological research on social stratification has conceived of the connection between social origins and educational and socioeconomic achievement in largely environmental terms. In the standard view the zero order positive association between individual educational achievement and socioeconomic status of the family of origin (measured, for example, by father's occupational status and parental education) is largely due to the greater availability of financial, cultural, and social resources (including network ties and positive role models) enhancing achievement of offspring from more privileged families. The zero order association between status of the family of origin and educational, occupational and income destinations of the offspring is then conceptualized as a (reverse) summary indicator of openness of the social structure, with a strong origin-destination association indicating fewer opportunities for mobility (greater system closure) and a weaker origin-destination association indicating more opportunities for mobility (greater system openness). A closely related concept (the intergenerational elasticity of income) is used by economists as a measure of intergenerational income mobility (e.g., Solon 2008). Much comparative research has used the association between origin and destination status as an inverse indication of the relative openness of the social stratification system in investigating social-developmental trends or differences among societies based on different political systems or types of welfare state (Beller and Hout 2006; Breen and Jonsson 2005; Ganzeboom, Treiman and Ultee 1991).

An equally venerable tradition of research, albeit one that has been less dominant in the social sciences, has emphasized the role of biological, hereditary mechanisms in educational and socioeconomic achievement (Galton 1869). This nature-oriented approach to social stratification was adopted by a relatively small group of social and behavioral scientists in association with the interdisciplinary field of quantitative genetics (Eckland 1967, 1979; Jencks 1972; Scarr-Salapatek 1971; Taubman 1976). In this quantitative genetic approach the association between origin and destination status is conceptualized as possibly due in part to the transmission from parents to offspring of genes that enhance or detract from educational and socioeconomic achievement, in addition to the environmental factors that are emphasized in mainstream stratification research.¹ The quantitative genetic approach recognizes that individual educational attainment may be associated with father's or mother's educational attainment in part because individuals inherit from both parents genes that affect their aptitude for educational pursuits.² As Eckland (1967) recognized a long time ago in sociology, the possible role of genes in the intergenerational association of status has far-reaching consequences for understanding mobility processes, a crucial point that will be further discussed later.

The conceptual framework and statistical methodology of quantitative genetics originated as part of the modern synthesis of Darwinian evolutionism and Mendelian genetics during the first half of the twentieth century (Fisher 1918;

¹We use the term "quantitative genetic" in preference to the alternative, commonly used term "behavior genetic", as the former better conveys the generality of the approach, which applies equally well to traits, such as body-mass index or a susceptibility to asthma, which do not constitute "behaviors" in the usual sense of the term.

²Early social mobility research pragmatically focused on father's occupational prestige and educational attainment because of the large proportion of mothers without tertiary education or occupation outside the home, a situation that is no longer common (Buchmann and DiPrete 2006). The quantitative genetic perspective inherently encourages a symmetrical treatment of parental influences, as both parents transmit to their offspring genes that can affect socioeconomic success.

Wright 1921a, 1921b, 1934) and were further developed to a high degree of sophistication in such fields as plant and animal breeding, before being shown to be applicable to human physical and psychological traits (Jinks and Fulker 1970; Martin and Eaves 1977). Human quantitative genetic research requires data on twins, adoptees, and other relatives that permit estimation of genetic and environmental components of the variance in the trait of interest, in contrast with the random samples of unrelated individuals typically used in social science research. Research in the quantitative genetic tradition has uncovered two major patterns concerning the role of genes in the etiology of human traits and behaviors. First, many human traits exhibit large effects of genes, and relatively smaller effects of environmental factors (Freese 2008; Turkheimer 2000). Second, for the well-studied trait of intelligence (measured as IQ), there is a systematic pattern of change in the relative roles of genes and the family environment over the life course: in childhood the family environment and genes are both important, but by early adulthood the role of the family environment has faded out and genes become the principal or only systematic determinant of the trait (Bouchard and McGue 2003; Petrill and Wilkerson 2000).

The patterns of a predominant role of genes and the fading role of the family environment over the life course uncovered by quantitative genetic research of human traits related to cognition would seem to have important implications for understanding social stratification and mobility processes. To the extent that educational and socioeconomic achievement of individuals are substantially determined by their genes rather than environmental processes in the family of origin, this needs to be taken into account in interpreting social mobility processes. For example, as pointed out in a pioneering paper by Eckland (1967), a significant positive intergenerational association on a status dimension such as occupational prestige may not be taken as direct evidence of the presence of environmental mechanisms of social reproduction. Sons and daughters from more prestigious origins may disproportionately end up in more prestigious destinations simply because they are more likely than offspring from less prestigious origins to inherit genes that allow entry into more prestigious destinations. Properly interpreting the intergenerational association of status thus requires separating the respective roles of genes and environment in the association. Furthermore, if measures of educational and socioeconomic success behave in the same way as IQ, with the role of the family environment disappearing by early adulthood, some fundamental assumptions on the role of the family of origin in individual achievement may need to be revised. In particular the part of the intergenerational association on a measure of socioeconomic success that is due to genetic transmission may have to be interpreted differently than the part of association due to environmental characteristics of the family of origin. In a nutshell, greater heritability (proportion of variance in achievement due to genetic variation) is indicative of greater opportunities for mobility, while greater environmentality (proportion of variance due to shared environmental factors) is indicative of less opportunities and stronger social reproduction of advantage (Guo and Stearn 2002; Jencks and Tach 2006; Nielsen 2006).

In this paper we focus on one measure of educational achievement, educational attainment measured as highest degree earned. We use data from the AddHealth study on 1,576 pairs of siblings aged 24 to 32 living in the same household at the time of the initial survey, classified into six types according to relatedness: monozygotic twins (MZ), dizygotic twins (DZ), full siblings (FS), half siblings (HS), cousins (CO) and non related siblings (NR). We use quantitative genetic modeling techniques based on structural equations models (SEM) to estimate the relative roles of genes and the family environment in educational attainment. We find that while genes play a role in attainment, the family environment remains a surprisingly strong determinant of educational attainment among these young adults. The role

of genes in educational attainment is relatively weaker (23 percent of the variance in attainment) and the role of the shared family environment stronger (41 percent of the variance for twins and 30 percent of the variance for nontwin siblings) than is typically found for cognitive outcomes (such as IQ) at the same young adult stage in the life course. This pattern of persisting influence of the family environment, which is at variance with the usual findings for intelligence, cannot be explained as a result of assortative mating of parents of the siblings or direct reciprocal effects of twins on each other's achievement.

In the next section we develop theoretical expectations concerning the genetic and environmental determinants of educational attainment. Later sections present the data and describe the SEM methodology. We then present results of three kinds of analysis: (1) a classic quantitative genetic decomposition of years of education, that reveals a stronger shared environmental component than is typical for achievement measures at this stage in the life course; (2) alternative variance decompositions taking into account of the strong degree of assortative mating on education in these data ($r = .629$ for parental education); (3) alternative variance decomposition for twins allowing for direct effects of education of the twins on each other. We conclude with a discussion of the results and their broader implications for social stratification research.

2 Environment, Genes and Success

Mainstream social stratification and social mobility research has traditionally looked at socioeconomic achievement as a series of measures ordered over the life course of the individual – from characteristics of the family of origin such as father's education and occupational status, through measures reflecting individual aptitudes (such as IQ scores) and relevant environmental influences (such as peers expectations), through educational attainment, to status of the destination occupation measured as occupational prestige and income. Earlier variables in this long causal chain are viewed as predictive of later ones (Blau and Duncan 1967; Duncan and Hodge 1963). A crucial pattern discovered by Blau and Duncan (1967) is that education constitutes the main factor in both upward mobility of individuals and the reproduction of status from generation to generation, so that when educational attainment is statistically controlled, little association between status of origin and destination remains (Hout and DiPrete 2006). Blau and Duncan (1967) also found that socioeconomic success through education seems to depend to a considerable extent on characteristics of individuals that are both unmeasured and independent of social origins. As Hout and DiPrete write (2006: 6): "The variation in education that is independent of social origins contributes more variance to destination than the portion of variation in education that comes from origins does." While intelligence may constitute part of these unmeasured individual factors, many researchers suspect that personality traits, interest patterns and motivational factors may also play a role (Carey 2003).

Social stratification researchers in sociology and economics sometimes explicitly recognize that genetics can play a role in the association between family background variables and individual achievement measures (e.g., Duncan, Featherman and Duncan 1972; Solon 2008), but could not resolve the roles of genes and environment using the random samples of unrelated individuals typically used in social research. Quantitative genetic researchers use special genetically informative designs – such as comparisons of identical and fraternal twins, or data on adopted children – to distinguish three components of the total variance in a trait, called a *phenotype* (such as height, score on an IQ test or years of education).

Heritability is the proportion of variance attributed to all causes of genetic

variability.³ The *shared environment* is the proportion of variance attributed to all environmental influences that are shared by siblings but vary between families. The shared environment component is potentially of paramount interest to sociologists, as it includes effects on the phenotype of variables such as social class, quality of local school systems, or ethnic culture. In the context of social stratification the shared environment measures the strength of “social ascription” (Conley 2008), or inequality associated with privilege. As Behrman and Taubman write: “[t]he share of the observed variation in schooling that is attributable to across-family variability in environment provides a measure of inequality in schooling opportunity” (1989: 1426, cited in Miller, Mulvey and Martin 2001: 214; see also Guo and Stearns 2002; Nielsen 2006; Jencks and Tach 2006; Visscher, Hill and Wray 2008). The shared environment also has a direct policy interpretation, as it reflects the potential effect on educational attainment of raising the quality of the most disadvantaged family environment to the level of the most advantaged ones. It thus represents an upper bound on improvement in the trait achievable by policy intervention *within the existing range of environmental variation* (Rowe 1994). Finally the *nonshared environment* represents the combined effects of factors that tend to make siblings different on the trait, including random measurement errors.

Quantitative genetic methods have been applied to a great variety of human traits, with substantial effects of genes found in many cases (Freese 2008; Freese and Shostak 2009). With some interesting exceptions (Nielsen 2008) the findings of many studies can be fairly summarized into what Turkheimer (2000: 160) has called “three laws of behavior genetics”:

- (1) “All human behavioral traits are heritable” – i.e., heritability tends to be substantial for most traits.
- (2) “The effect of being raised in the same family is smaller than the effect of the genes” – i.e., the shared environment is typically smaller than heritability.
- (3) “A substantial portion of the variation in complex human behavioral traits is not accounted for by the effects of genes or families” – i.e., the nonshared environment is substantial.

Quantitative genetic research on intelligence has uncovered an additional pattern that may be of relevance to social stratification research more generally. A systematic pattern of change in the components of IQ variance over the life course has emerged. Heritability of IQ is typically found to be about 45 percent in childhood, with the shared environment at 35 percent. In adulthood heritability rises to over 80 percent, and the role of the shared environment vanishes. Throughout the life course the nonshared environment represents some 20 percent of the IQ variance (Jensen 1998; Boomsma, Busjahn and Peltonen 2002; Bouchard and McGue 2003). Thus, at some point in late adolescence or early adulthood, shared family environmental factors cease to be an importance component of the variance in cognitive ability. A similar trend may well characterize measures of educational achievement that may be assumed to have a strong cognitive component, such as school grades or SAT and GRE scores (e.g., Frey and Detterman 2004; Koenig, Frey and Detterman 2008). Given the effects of these cognitive measures on continuation at different stages of an educational career, one might further expect the pattern of fading shared environmentality to be reflected in educational attainment even when the latter is measured non-cognitively as years of education or highest degree earned.

Quantitative genetic research on educational achievement is relatively limited compared to the voluminous research on intelligence (Carey 2003). However research in a variety of fields has focused on a number of traits related to educational

³Quantitative genetic researchers distinguish between *narrow sense heritability*, which is the proportion of variance due to the additive effects of genes, and *broad sense heritability*, which includes effects of dominance and epistasis. However, we do not distinguish between these two forms of heritability in this paper.

achievement. This research, which is by necessity based on adult subjects when achievement is measured as years of education completed or highest degree, typically finds substantial heritability, a pattern consistent with the intelligence research (Baker et al. 1996; Behrman and Taubman 1989; Heath et al. 1985; Lichtenstein, Pedersen and McClearn 1992; Miller, Mulvey and Martin 2001; Tambs et al. 1989; Vogler and Fulker 1983). In contrast with intelligence research, however, there are also reports of substantial shared environment effects (e.g., Heath et al. 1985), although this result may be in part an artefact of the high degree of assortative mating on educational achievement (Miller, Mulvey and Martin 2001).

There is a variety of ways to measure educational achievement, which may vary in their dependence on genetic and environmental factors. In some cases achievement is measured as grade point average (GPA), although little is known of the quantitative genetics of GPA (Carey 2003; but see Nielsen 2006). In other situations educational achievement is measured as the score on a standardized test bearing on materials related to school curriculum, or the score on a standardized academic aptitudes test (such as the SAT or the GRE). In still other cases educational achievement is measured as the highest degree earned (high school, some college, college, postgraduate degree), or simply as years of education completed.

The quantitative genetic architectures of these various measures of educational achievement would be expected to differ according to at least two dimensions: (1) The extent to which the measure is based on cognitive tests, and thus incorporates a cognitive component related to intelligence; (2) The extent to which the achievement measure is directly affected by resources (including financial) available to the family. Measures of achievement based on tests (such as the SAT) or aggregated from collections of tests (such as GPA) would be expected to have a strong cognitive component, and to be similar in that way to measures of intelligence (IQ tests), a conjecture consistent with the strong and well documented association between intelligence and school achievement measured in that way (Deary et al. 2007; Frey and Detterman 2004; Koenig, Frey and Detterman 2008). In the context of stratification processes, one common feature of these test-based cognitive measures of educational achievement and measures of intelligence is that families of the individual cannot by their resources or actions do much to *directly* affect the score of the individual on the tests. Even if one recognizes that families have available different amounts of resources to pay for test preparations and other activities that may raise scores, the actual effectiveness of test preparation in raising scores remains controversial (Briggs 2001; Buchmann, Condron and Roscigno 2010). For the most part family members cannot stand behind their test-taker and help them by whispering the correct answers to questions. One would expect that the quantitative genetic patterns for measures of educational achievement involving a large cognitive component, such as GPA or SAT scores or the entrance examination for the *École Normale Supérieure*, might be similar to that for intelligence. One might therefore conjecture that these cognitively-loaded measures of educational achievement will behave similarly to intelligence, with the shared environmental component declining by early adulthood and genes becoming the only systematic source of variation.

Educational achievement measures reflecting transitions between levels of education (Mare 1980, 1981), such as college entrance or college graduation, are more likely to be directly affected by resources and other characteristics and activities of the family. Years of education completed and highest degree earned, which are overall summaries of a number of such transitions, would also be more strongly dependent on the family. As Jencks and Tach (2006) have observed, a college education, at least in systems such as the U.S. where higher education may represent a considerable financial burden, entails the writing by someone of a series of large checks. One might therefore predict that the impact of the family, measured as the

shared environmental component, will remain larger for measures of educational achievement based on transition to college and graduation from college than it is for more purely cognitive measures. This conjecture is consistent with the finding that parental income while a student is in high school has a significant direct effect on college attendance, but does not affect other outcomes after controlling for common causes (Mayer 1997). Generally speaking, one would expect the role of the shared environment to fade more slowly for these measures over the life course than for cognitive measures.

In the remainder of the paper we investigate the quantitative genetic decomposition of educational attainment measured as highest degree earned for samples of young adult siblings. Based on the discussion of this section, we expect to find that the shared environment component of the quantitative genetic decomposition of the variance in attainment estimated for these young adults will remain significant and substantial, unlike a purely cognitive measure such as IQ for which the shared environment component has typically become zero in the young adult age range.

3 Data and Methods

3.1 Data

We use data from the National Longitudinal Study of Adolescent Health (Add Health), a longitudinal study of a nationally representative sample of adolescents in grades 7–12 in the United States during the 1994–95 school year. The Add Health cohort has been followed into young adulthood with four in-home interviews, the most recent (wave 4) in 2008, when the sample was aged 24–32 (Harris 2009).⁴

The measure of educational attainment for respondents is based on the wave 4 question “What is the highest level of education you have achieved to date?” (Add Health variable `h4ed2`). Responses were recoded from 13 to 11 categories by combining academic and professional graduate training (into categories 10 and 11, respectively) resulting in the scale `w4edu`: 1 = 8th grade or less; 2 = some high school; 3 = high school graduate; 4 = some vocational/technical training (after high school); 5 = completed vocational/technical training (after high school); 6 = some college; 7 = completed college (bachelor’s degree); 8 = some graduate school; 9 = completed a master’s degree; 10 = some graduate training beyond a master’s degree or some post baccalaureate professional education (e.g., law school, med school, nurse); 11 = completed a doctoral degree or completed post baccalaureate professional education (e.g., law school, med school, nurse).

The measure of parental education used to estimate assortative mating was based on parental questionnaire items “How far did you go in school?” (`pa12`) and “How far did your current (spouse/partner) go in school?” (`pb8`). The original 10 categories were recoded resulting in the scale: 0 = less than 8th grade; 1 = less than HS but more than 8th grade; 2 = HS graduate or GED; 3 = some education beyond HS (vocational or College); 4 = graduated from a college or university; 5 = professional training beyond a 4-year college or university. Where data from parents were missing, wave 1 items “How far in school did he/she go?” asked of students about their parents (`h1rm1` and `h1rf1`) and recoded the same way were used.

Recent youth cohorts in the U.S. show a female advantage in educational attainment (Buchmann and DiPrete 2006). There is also a possibility that younger students in the study are more likely to still be in school, so the highest degree they report is less than the one they will eventually achieve. Table 1 shows the regression of `w4edu` on age and female gender. The nonsignificant regression coefficient for age suggests there is no censoring of further educational achievement for younger

⁴Add Health codebooks are available online at <http://www.cpc.unc.edu/projects/addhealth/codebooks>.

Table 1: Regression of Wave 4 Educational Attainment (`w4edu`) on Age (`w1age`) and Female Gender (`female`) (N=13,622 subjects)

Variable	Coef.	t	P> t	[95% CI]	
Age	-.014	-1.45	0.147	-.034	.005
Female Gender	.525	15.24	0.000	.457	.592
Intercept	5.486	35.09	0.000	5.179	5.792
R-squared	0.017				

respondents. However the significant positive coefficient on the female indicator reveals a substantial female advantage in educational attainment. To correct for the female advantage and simplify the analysis, we adjust the educational attainment variable for sex only (but not for age) using the formula:

$$w4edut = w4edu - 0.525 \times female$$

where `w4edut` is the transformed measure of educational attainment, `w4edu` is the original variable, and `female` is an indicator variable with value 1 for female and 0 for male.

In the first wave of AddHealth siblings living in the same household were identified, and if necessary, a sibling was added to the sample to complete a pair. Pairs were classified as monozygotic twins (MZ), dizygotic twins (DZ), full siblings (FS), half siblings (HS), cousins (CO), and nonrelated (NR) (on determination of zygosity see Rowe and Jacobson 1998 and Harris et al. 2006). Numbers of pairs of each type actually used in the analysis are shown in Table 2. All sibling pairs are used, even though data on an individual may be repeated when an individual is member of more than one pair (see discussion of this point in Eaves et al. 1999: 67).

3.2 Methods

The methodology of quantitative genetics originates in pioneering contributions by Fisher (1918), Wright (1921a, 1921b, 1934) and Jinks and Fulker (1970). Neale and Cardon (1992) and Neale and Maes (2004) provide extensive treatments. Cherny (2008), Kohler, Behrman and Schnittker (2011), and Medland and Hatemi (2009) provide overviews oriented to social scientists. The method is based on the decomposition of the covariance matrix of a trait across relatives (i.e., in this paper the covariance of educational attainment across siblings who vary in their relatedness) into genetic and environmental components. The basic ACE model is shown in Figure 1. The symbols P_1 and P_2 enclosed in squares represent measures of the phenotype (in this case educational attainment) for the first and second sibling in a pair, respectively. The model presents P_1 and P_2 as linear functions of three latent (unmeasured) factors enclosed in circles. A represents all additive effects of genes on the value of the trait. The model assumes that the values A_1 and A_2 of A for the siblings are correlated by a quantity k that corresponds to the degree of relatedness, the (average) proportion of genes shared by the siblings by common descent. For MZ twins, whose genomes are identical, $k = 1.0$. Assuming random mating $k = 0.5$ for DZ twins and full siblings; $k = 0.25$ for half siblings; and $k = 0.125$ for cousins. Finally, $k = 0$ for nonrelated siblings.

Latent variable C stands for the shared environment, representing all aspects of the family and larger environment (such as the neighborhood) that are shared by siblings and tend to make them similar on the trait. C would include social class, neighborhood quality, and various cultural characteristics of the family environment

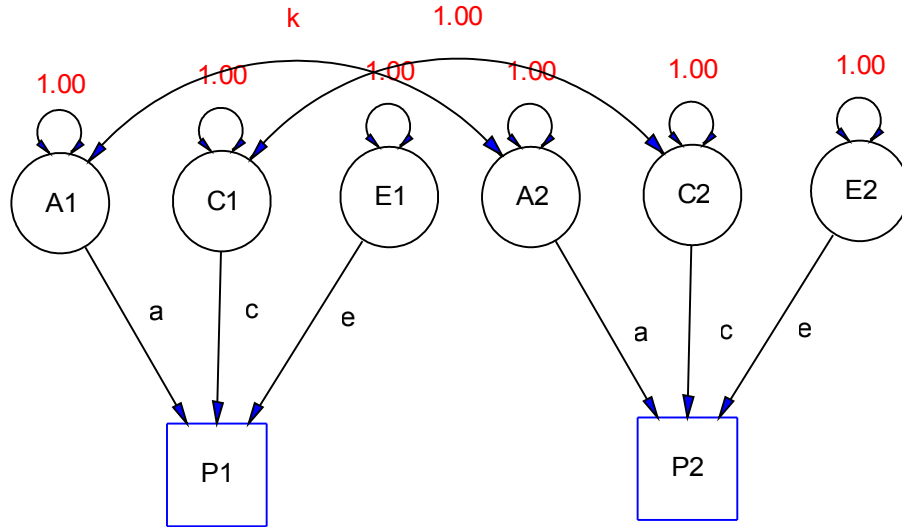


Figure 1: Basic ACE Model of Quantitative Genetics.

that affect siblings in the same ways. The values of C are assumed to be the same for both siblings (correlation = 1.0). Finally latent variable E represents the nonshared environment consisting of all aspects of the environment and life experiences that differ across siblings and tend to make them *different* on the trait. E would include effects of a parental preference for one sibling, a childhood disease, an encounter with an inspiring mentor, or peer influences that affect the siblings differently, as well as measurement error. The values of E are assumed to be uncorrelated across siblings.

Using the tracing rules of path analysis, the expected covariance matrix Σ of the trait among siblings in a pair is given by

$$\Sigma = \begin{pmatrix} a^2 + c^2 + e^2 & ka^2 + c^2 \\ ka^2 + c^2 & a^2 + c^2 + e^2 \end{pmatrix}$$

where a^2 is the variance attributable to the genetic influences A , c^2 is the variance attributable to environmental influences C shared by a pair of twins, and e^2 is the variance attributable to environmental influences E unique to the individual. If these three quantities are standardized to the phenotypic variance (so they add up to 1.0), then a^2 represents *heritability*, the proportion of variance due to genetic influences, and c^2 and e^2 represent *shared* and *nonshared environmentality*, respectively.

There is a separate expected covariance matrix for each type of siblings, with the fixed parameter k set to the corresponding degree of relatedness. For analyses based on data for twins only there are two expected covariance matrices, one with $k = 1$ for MZ twins and one with $k = 0.5$ for DZ twins. For analyses based on all sibling types there are six expected covariance matrices. The parameters of the model are estimated by maximizing the log likelihood ML of the data across all types of sibling pairs according to the formula

$$ML = df(\ln |\Sigma| - \ln |\mathbf{S}| + (\text{tr}(\mathbf{S}\Sigma^{-1})) - p)$$

where \mathbf{S} is the observed covariance matrix, Σ is the expected covariance matrix, $\text{tr}(\mathbf{A})$ indicates the trace of matrix \mathbf{A} , $|\mathbf{A}|$ indicates the determinant of matrix \mathbf{A} , p is the number of observed variables ($p = 2$ in this case), and df is one less than the sample size (number of pairs). We estimate the models using the program Mx (Neale et al. 2003).

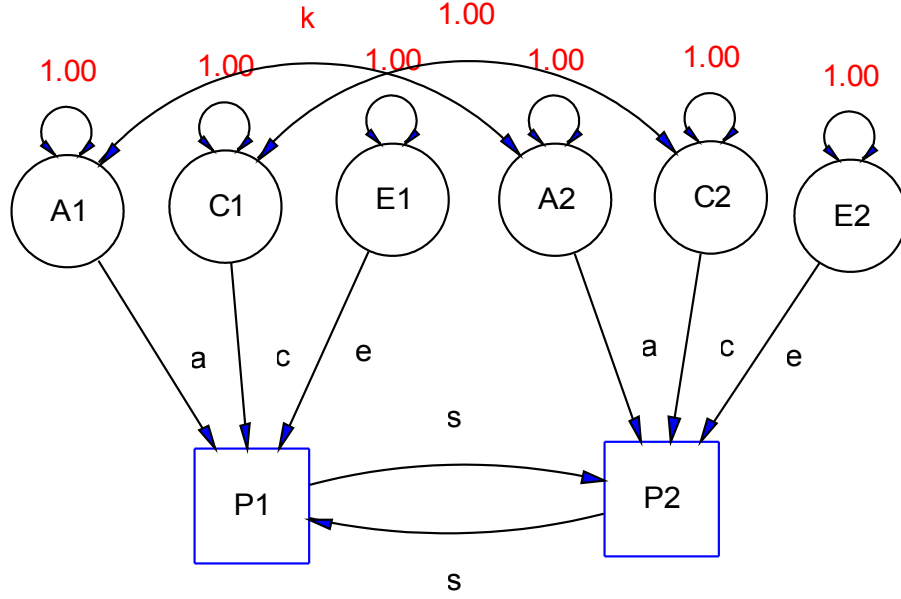


Figure 2: ACE Model with Direct Sibling Interaction.

In some analyses we use an extension of the ACE model shown in Figure 2 that allows for direct phenotypic interactions between siblings. In addition to the latent components A , C , and E , that have the same interpretations as in Figure 1, the model specifies that the phenotypic value of each sibling directly and linearly affects the value of the other. This direct phenotypic effect is denoted by the coefficient s , with $|s|$ assumed less than 1 (so the system tends to a finite value). A positive value of s corresponds to a cooperative interaction such that an increase in the value of P in one sibling enhances P in the other; a negative value of s corresponds to a competitive interaction such that an increase in the value of P in one sibling results in a decrease in P in the other. It can be shown that the expected covariance matrix Σ of the trait among siblings in a pair is given by

$$\Sigma = (\mathbf{I} - \mathbf{B})^{-1} \times \begin{pmatrix} a^2 + c^2 + e^2 & ka^2 + c^2 \\ ka^2 + c^2 & a^2 + c^2 + e^2 \end{pmatrix} \times ((\mathbf{I} - \mathbf{B})^{-1})'$$

where

$$B = \begin{pmatrix} 0 & s \\ s & 0 \end{pmatrix}$$

and s indicates the direct effect of one sibling on the other (Neale and Maes 2004).

4 Results

Table 2 shows the covariance and correlation matrices for the six groups of siblings. While the structural equations models discussed below are estimated from the covariances, we first look at the patterns of correlations, as these are easier to interpret. The correlation of educational achievement for MZ twins (.636) is larger than the correlation for DZ twins (.559), suggesting a role of genes in educational attainment. However the role of genes appears to be small. A classical formula to estimate heritability from twins data simply doubles the difference between the correlations for MZ and DZ twins: $a^2 = 2 \times (r_{MZ} - r_{DZ}) = 2 \times (.636 - .559) = .154$, suggesting that only about 15 percent of the variance in educational attainment is due to genes. The correlation for DZ twins (.559) is unexpectedly larger than the

Table 2: Covariances and Correlations of Gender-Corrected Educational Attainment (*w4educ*) by Siblings Type

Siblings Type	N Pairs	Covariance Matrix		Correlation Matrix	
MZ – Monozygotic Twins	200	4.54391		1.0	
		2.69911	3.96422	0.6360	1.0
DZ – Dizygotic Twins	324	4.31029		1.0	
		2.3685	4.16843	0.5588	1.0
FS – Full Siblings	639	4.18372		1.0	
		1.79389	3.8643	0.4461	1.0
HS – Half Siblings	213	4.10514		1.0	
		.757862	3.63197	0.1963	1.0
CO – Cousins	68	3.26292		1.0	
		1.39484	4.65565	0.3579	1.0
NR – Non Relateds	132	3.41589		1.0	
		1.08983	3.86927	0.2998	1.0
Total	1576				

correlation for full siblings (.446); on the basis of genes alone one would expect these correlations to be the same, since DZ twins and full siblings are genetically related to the same degree, sharing an average 50 percent of their genes. The larger correlation for DZ twins may therefore indicate a “contemporaneity” effect generating greater similarity of the shared environments for twins, who are the same age, than for ordinary siblings, who are born at different times; we will formally investigate this possibility later. While the correlation for half siblings (.193) is about half that for full siblings (.446), a pattern expected on the basis of genetic relatedness alone, the correlations for cousins (.358), who share only one eighth of their genes, and for non-related siblings (.300), who share none, are unexpectedly high. The higher than genetically-expected correlations for cousins and non-related siblings, and the fact that correlations for DZ twins and full siblings are both larger than half the correlation for MZ twins, are clues to a substantial effect of the shared environment on the educational attainment of siblings.

The utility of the structural equations approach is that we can simultaneously test the statistical significance of apparent discrepancies from the theoretically expected pattern found in Table 2. Results from fitting the SEM models are shown in Table 3. The first panel of Table 3 shows the results of fitting the classic ACE model and its submodels to the data for twins only. Line 1 of the table shows the fit statistics for the full ACE model including genetic (a^2), shared environmental (c^2) and unshared environmental (e^2) components. The components are standardized so that they add up to one. The largest estimated component of the variance in educational attainment (.484) corresponds to c^2 , the shared environment component that summarizes environmental influences that tend to make siblings similar on the trait. The next largest component (.366) corresponds to the unshared component e^2 , which includes any measurement error as well as sibling-specific environmental influences that tend to make the twins different from each other. The heritability a^2 , which estimates the proportion of the variance in attainment due to genes, is relatively small at .150, close to the .154 found earlier using the classic formula for heritability. The ACE model as a whole has a χ^2 of 1.685 for three degrees of freedom, corresponding to a nonsignificant P-value of .640 which represents an

Table 3: ACE Models of Educational Attainment – Estimates for Twins Only and for All Siblings Types

Model	Parameter Estimates			Fit Statistics				
	a^2	c^2	e^2	χ^2	df	p	AIC	RMSEA
Twins Only (2 groups, N = 524 pairs)								
1. ACE	.150	.484	.366	1.685	3	.640	-4.315	.007
2. CE	— ^a	.588	.412	3.917	4	.417	-4.083	.025
3. AE	.675	— ^a	.325	30.081	4	.000	22.081	.132
4. E	— ^a	— ^a	1.0	225.047	5	.000	215.047	.417
All Siblings Types (6 groups, N = 1576 pairs)								
5. ACE	.365	.272	.364	22.804	15	.088	-7.196	.032
6. CE	— ^a	.448	.552	52.108	16	.000	20.108	.090
7. AE	.696	— ^a	.304	64.591	16	.000	32.591	.112
8. E	— ^a	— ^a	1.0	404.000	17	.000	370.000	.259

Note: Parameters are a^2 , additive genetic; c^2 , shared environment; e^2 , nonshared environment.

^a Parameter fixed to value zero.

excellent fit.⁵ The adequacy of the ACE model is further confirmed by the value of the root mean squared error of approximation (RMSEA) which, at .007, indicates a very good fit.⁶

The significance of each component of the variance in educational attainment among twins is further evaluated in the other tests in the first panel of Table 3. Lines 2 to 4 show the consequences for model fit of dropping (i.e., forcing to zero) various components of the ACE model. Line 2 shows that the model CE that excludes the genetic component a^2 corresponds to a χ^2 of 3.917 with 4 df. As the CE model of Line 2 is nested within the ACE model of Line 1, the significance of a^2 can be tested by subtracting the χ^2 s of the ACE model from the CE model, and corresponding degrees of freedom, as $3.917 - 1.685 = 2.232$ for $4 - 3 = 1$ df, which is nonsignificant ($p = .135$). We therefore cannot reject the possibility that $a^2 = 0$ and there is no genetic influence at all on the educational attainment of twins. By contrast, setting the shared environment c^2 to zero in the model AE (on Line 3) produces a substantial deterioration in fit: the difference in χ^2 is $30.081 - 1.685 = 28.397$ with $4 - 3 = 1$ df, which is highly significant ($p = .000$). Thus the hypothesis that there is no shared environment effect ($c^2 = 0$) can be rejected. Finally, as shown on Line 4 of the table, the model E with both shared environment and genetic component set to zero can also be rejected as it corresponds to a massively significant increase in χ^2 compared to the ACE model.

While we did expect to find a higher value of the shared environment component c^2 , representing the impact of family and other shared environmental influences on educational attainment, in comparison with other more cognitive measures of

⁵The degrees of freedom are calculated as follows. The data for twins, with three statistics (two variances and one covariance) for each of the two groups (MZ and DZ), provide $2 \times 3 = 6$ df. The ACE model estimates three parameters (a^2 , c^2 and e^2), using up 3 df. The test of the full ACE model (Line 1 in Table 3) is thus based on $6 - 3 = 3$ df. Constraining a parameter to be zero frees up one degree of freedom, so for example the model CE on Line 2 of the table, which constrains a^2 to be zero, has $3 + 1 = 4$ df, one df more than the full ACE model of Line 1. Constraining two parameters to have the same value also yields an additional df. Similar calculations apply to the other models and submodels estimated in the paper. In analyses using all 6 groups of siblings the data provide $6 \times 3 = 18$ df.

⁶Values of RMSEA below .10 are deemed to indicate a good fit, and values below .05 a very good fit (Bollen 1989).

achievement, the lopsided role of the shared environment in explaining 48% of the variance in attainment in the twins data, together with the small heritability of 15%, is an unusual finding, representing as it does a striking exception to Turkheimer's (2000) "second law" that the effect of genes is greater than the effect of the family. As discussed earlier, the typical pattern with cognitive measures is one of fading of the shared environment in young adulthood, with c^2 trending to zero over the life course and heritability a^2 becoming the major component of variance. We take advantage of the availability of data on other types of siblings to estimate the same models with all six groups of siblings.

Results for the six groups of siblings are shown in the second panel of Table 3. The full ACE model on Line 5, corresponding to a χ^2 of 22.804 with 15 df, still represents an acceptable fit to the data ($p = .088 > .05$). (Part of the decrease in fit compared to the twins only data is undoubtedly due to the larger number of pairs.) The genetic component a^2 at .365 is now substantially larger than the shared environmental component c^2 at .272. This result is more in line with the typical pattern for human traits, as summarized in Turkheimer's (2000) "second law" that the effect of the genes tends to be larger than the effect of the shared environment. The unshared environment e^2 is practically unchanged at .364 (compared to .366 for twins). Lines 6 to 8 in Table 3 test the consequences of dropping in turn a^2 (model CE), c^2 (AE), or both a^2 and c^2 (E). Each of these submodels now corresponds to a significant increase in the chi-square of the model relative to the ACE model and can thus be rejected. For example, comparing the model CE that drops the a^2 component (Line 6) with the full ACE model (Line 5) produces a difference in χ^2 of $52.108 - 22.804 = 29.304$ for $16 - 15 = 1$ df, which is highly significant ($p = .000$).

The conclusions we reached on the basis of comparing the χ^2 s of different models are confirmed in a comparison of the additional fit statistics also provided in Table 3: Akaike's information criterion (AIC) and the RMSEA. The AIC can be used to compare the fit of different models, even when they are not nested. A more negative value of AIC indicates a better model. In the upper panel of the table (twins data only) the best-fitting model is ACE (AIC = -4.315), closely followed by the CE model (AIC = -4.083) which, as we saw, is not significantly different from ACE in terms of fit. ACE and CE, with RMSEA of .007 and .025, respectively, are also the only two models in the upper panel that satisfy the RMSEA $< .05$ criterion for a very good fit. Among models in the lower panel of Table 3 ACE is clearly the best model as indicated by the lowest AIC value (-7.196); it is also the only model providing a very good fit (RMSEA = $.032 < .05$).

The ACE models estimated for twins only and for all siblings types (including twins) in Table 3 show substantial differences. The ACE model for twins only shows lower heritability and a much greater role of the shared environment ($a^2 = .150$, $c^2 = .484$) compared to the ACE model estimated for all types of siblings ($a^2 = .365$, $c^2 = .272$). To further explore these differences we estimated models that allow the environmental components c^2 and e^2 of the variance in educational attainment to differ for twins and nontwin siblings. This is achieved by postulating a model in which the heritability a^2 , measuring genetic effects, has the same value for twins and nontwin siblings, in conformity with genetic theory. However the model allows the shared environment component to have different values, denoted c^2_{twi} for twins and c^2_{sib} for nontwin siblings. Substantively, this is a way to model the fact that twins always are the same ages, while nontwin siblings are born at different times. Being exact contemporaries, twins may experience more similar environments than nontwin siblings. As the model continues to constrain the (total) expected variance in educational attainment to be the same for twins and nontwin siblings, it follows that the nonshared environment components also have different values, denoted e^2_{twi} for twins and e^2_{sib} for nontwin siblings (see Chipuer, Rovine and Plomin 1990). Estimates for this special twins environments model and submodels are shown in

Table 4: ACE Models of Educational Attainment with Special Twins Environments (N = 1576 pairs)

Model	Parameter Estimates					Fit Statistics			
	a^2	c_{twi}^2	e_{twi}^2	c_{sib}^2	e_{sib}^2	$\chi^2(df)$	p	AIC	RMSEA
1. ACE	.228	.411	.361	.298	.474	17.514(14)	.230	-10.485	.013
2. ACE ^a	.365	.272	.364	.272	.364	22.804(15)	.088	-7.196	.023
3. CE	— ^b	.573	.427	.380	.620	25.252(15)	.047	-4.748	.023
4. AE	.696	— ^b	.304	— ^b	.304	64.591(16)	.000	32.591	.093
5. E	— ^b	— ^b	1.0	— ^b	1.0	404.000(17)	.000	370.000	.224

Note: Parameters are a^2 , additive genetic; c_{twi}^2 , twins shared environment; e_{twi}^2 , twins nonshared environment; c_{sib}^2 , nontwin siblings shared environment; e_{sib}^2 , nontwin siblings nonshared environment.

^a Model constrained so that $c_{\text{twi}}^2 = c_{\text{sib}}^2$ and $e_{\text{twi}}^2 = e_{\text{sib}}^2$.

^b Parameter fixed to value zero.

Table 4.

Results for the full ACE model with special twins environments are shown on Line 1 of Table 4. The χ^2 of 17.514 with 14 df represents a satisfactory fit to the data ($p = .230$). The RMSEA of .013 indicates that the fit is very good and the AIC of -10.485 is the lowest among all models estimated so far.⁷ The single estimate of a^2 , the heritability for both twins and nontwin siblings, is a moderate .228. The shared environment component c^2 is estimated at .411 for twins and .298 for nontwin siblings, consistent with the conjecture that environments of twins, who are the same age, are more similar (and therefore incorporate a greater proportion of shared experiences) as compared with nontwin siblings. The nonshared environments exhibit the opposite pattern, with values .361 and .474 for twins and nontwin siblings, respectively, consistent with the notion that nontwin siblings experience more dissimilar environments.

The model on Line 2 of Table 4, also a full ACE model, tests the equivalence of environmental components for twins and nontwin siblings by constraining the shared and nonshared environments to be the same for the two types of siblings. Because the sum of the variance components is constrained to be the same for twins and nontwin siblings, equality of the shared components entails equality of the nonshared components, so that the model on Line 2 specifies $c_{\text{twi}}^2 = c_{\text{sib}}^2$ and $e_{\text{twi}}^2 = e_{\text{sib}}^2$. Note that this is the same ACE model that was estimated earlier (on Line 5 of Table 3). As this constrained model is now viewed as nested within the unconstrained model on Line 1, we can compare the two models with respect to fit. The difference in χ^2 s is $22.804 - 17.514 = 5.289$ for $15 - 14 = 1$ df, indicating that the equality constraint on the environmental components entails a significant deterioration in fit ($p = .021 < .05$); in other words, there is a statistically significant difference between the environmental components for twins and nontwin siblings. The models CE, AE and E obtained by dropping a^2 , c^2 and both a^2 and c^2 , respectively, are shown on Line 3 to Line 5. All entail a significant deterioration in fit compared to the ACE model on Line 1 and can therefore be rejected. The full, unconstrained ACE model on Line 1 is therefore the favored model, a conclusion confirmed by the

⁷The ACE model on Line 1 constrains total expected variances for twins and nontwin siblings to be the same. Allowing total expected variances to differ improves the fit slightly ($\chi^2 = 15.702$ with 13 df, $p = .266$, AIC = -10.298, RMSEA = .028). As the total expected variances differ the (standardized) heritabilities for twins and nontwin siblings also differ: $a_{\text{twi}}^2 = .223$, $c_{\text{twi}}^2 = .429$, $e_{\text{twi}}^2 = 0.347$ and $a_{\text{sib}}^2 = .242$, $c_{\text{sib}}^2 = .285$ and $e_{\text{sib}}^2 = .473$. However the difference in fit with the model on Line 1 is nonsignificant ($\chi^2 = 17.514 - 15.702 = 1.812$ with $14 - 13 = 1$ df, $p = .178$).

Table 5: ML Confidence Intervals for ACE Model of Educational Attainment with Special Twins Environments (N = 1576 Pairs)

Parameter	Estimate	95% Confidence Interval	
		Lower bound	Upper bound
a^2	.228	.068	.390
c_{twi}^2	.411	.270	.538
e_{twi}^2	.361	.304	.428
c_{sib}^2	.298	.217	.376
e_{sib}^2	.474	.363	.589
$\chi^2(14)$	17.514	14.000	35.782
AIC	-10.485	-14.000	7.782
RMSEA	.013	.000	.058

Note: Parameters are a^2 , additive genetic; c_{twi}^2 , twins shared environment; e_{twi}^2 , twins nonshared environment; c_{sib}^2 , nontwin siblings shared environment; e_{sib}^2 , nontwin siblings nonshared environment.

minimal AIC value of -10.485 and the RMSEA of .013, corresponding to a very good fit.

Table 5 provides maximum-likelihood confidence intervals for the parameters of the ACE model with special twins environments and some of the fit statistics.⁸ All parameters in Table 5 have confidence intervals that exclude zero, consistent with previous likelihood ratio tests, although the lower bound for heritability a^2 comes closest (.068).

Estimated values of c^2 from the favored model of Table 5 are relatively large for twins ($c_{\text{twi}}^2 = .411$) as well as for nontwin siblings ($c_{\text{sib}}^2 = .298$) compared to typical estimates of the shared environment effect on educational achievement found in the literature. While there are good substantive reasons (mentioned earlier) for expecting a more substantial role of the shared environment for educational attainment as compared with more cognitive measures of educational achievement, a major alternative hypothesis needs to be addressed: that estimates of c^2 are inflated as a result of *assortative mating*, or *homogamy*, the tendency of people to find mates who are similar to themselves with respect to a given trait. Assortative mating on educational attainment, measured as the correlation of attainment between spouses, is known to be particularly high (Blossfeld and Timm 2003; Mare 1991; Rosenfeld 2008; Smits, Ultee and Lammers 1998). Assortative mating may affect quantitative genetic models of educational attainment in the following way: to the extent that educational attainment of parents is affected by their genes, a tendency of parents to find spouses with similar levels of educational attainment will cause the genes of the parents to be correlated, in the sense that they will share more of the genes affecting educational attainment than would be expected on the basis of random mating. This genetic consequence may be the outcome of purely social processes, such as college campuses functioning as “marriage markets” (Blossfeld and Timm 2003). As a result of the correlation between the genomes of the parents, the genomes of full siblings among their offspring will tend to be more highly correlated than the .5 value assumed under random mating. However, assortative mating does not affect the genomic correlation of MZ twins, as the latter share all their genes so the correlation

⁸See Neale et al. (2003) and Neale and Miller (1997). These confidence intervals do not depend on an asymptotic theory and need not be symmetric around the parameter estimate.

Table 6: Assortative Mating – Correlations of Educational Attainment of Father and Mother, by Siblings Type

Siblings Type	N Pairs	Combination	Correlation
MZ – Monozygotic Twins	118	Mother ₁ × Father ₁	.5740
		Mother ₂ × Father ₂	.5956
DZ – Dizygotic Twins	189	Mother ₁ × Father ₁	.7063
		Mother ₂ × Father ₂	.6789
FS – Full Siblings	454	Mother ₁ × Father ₁	.6211
		Mother ₂ × Father ₂	.6351
HS – Half Siblings	84	Mother ₁ × Father ₁	.6047
		Mother ₂ × Father ₂	.5470
CO – Cousins	22	Mother ₁ × Father ₁	.5100
		Mother ₂ × Father ₂	.2219
NR – Non Relateds	53	Mother ₁ × Father ₁	.5967
		Mother ₂ × Father ₂	.5868

Note: Overall degree of assortative mating estimated as $\hat{r} = 0.629$ (average of 10 correlations excluding cousins, weighted by number of pairs).

of their genomes is 1.0 no matter the degree of assortment of parents. Estimating the ACE model assuming random mating (i.e., that the genomic correlation = .5 for DZ twins and full siblings), as it depends on the difference between phenotypic correlations of MZ and DZ twins, will then produce an upward bias in the estimate of the shared environmental component c^2 , and a corresponding downward bias in the estimate of the heritability a^2 .

To investigate this issue we first estimate the degree of assortative mating as the correlation of educational attainment among the parents of the subjects. Table 6 shows the correlations of educational attainment of the father and mother of the subjects, separately by siblings type and for each sibling in a pair (denoted by subscripts 1 and 2).⁹ Correlations between parents for the 22 pairs of cousins with parental information are anomalously low (.510 and especially .222 for parents of the first and second sibling, respectively). Correlations for other sibling types vary between .706 and .547. Excluding cousins, the average correlation over all types and first and second sibling, weighted by the number of cases, is .629, which we adopt as an overall estimate of assortative mating. The large size of this correlation and its relative uniformity over siblings types are striking, especially given that for some types, e.g. half-siblings, the parents include at least one step-parent, rather than two biological parents of a pair of siblings.

Table 7 shows results for a modified ACE model applied to the twins only that relaxes the random mating assumption. Instead of being set at .5, the genomic correlation between DZ twins is iteratively estimated as part of the model, based on the formula given by Falconer (1989, Table 10.6 p. 178). The model assumes “phenotypic” assortative mating, i.e. that their resemblance on the trait is the direct cause of the tendency of spouses to mate together.¹⁰ This assumption allows derivation of the genotypic correlation between spouses due to assortment as $m = rh_0^2$ where m is the genotypic correlation, r is the phenotypic correlation between parents (estimated

⁹The mother and the father of each sibling in a pair may or may not refer to the same person, depending on siblings type.

¹⁰Falconer (1989, p. 176) notes that the assumption may not apply to humans, for whom the causes of spousal resemblance may be environmental as well as genetic. However, assuming phenotypic mating provides an upper bound for the effect of assortative mating in inflating c^2 .

Table 7: ACE Models With Assortative Mating Fit to the Educational Achievement Data – Twins Only (N = 524 Pairs)

Model	Parameter Estimates			Fit Statistics				
	a^2	c^2	e^2	χ^2	df	p	AIC	RMSEA
1. ACE	.167	.467	.366	1.685	3	.640	-4.315	.007
2. CE	— ^a	.588	.412	3.917	4	.417	-4.083	.025
3. AE	.678	— ^a	.322	7.662	4	.105	-.338	.060
4. E	— ^a	— ^a	1.0	225.047	5	.000	215.047	.417

Note: Parameters are a^2 , additive genetic; c^2 , shared environment; e^2 , nonshared environment.

^a Parameter fixed to value zero.

at .629) and h_0^2 is a provisional estimate of heritability. The genotypic correlation between full siblings (including DZ twins) is then calculated as $k = .5(1 + m)$ and updated values for h^2 and other variance components estimated, the model converging quickly to the values shown in Table 7 (see Chipuer, Rovine and Plomin 1990 for a “manual” version of this procedure). For the full ACE model on Line 1 of the table, a^2 is now .167 and c^2 .467, only slightly different from the corresponding estimates .150 and .484 based on the random mating assumption (upper panel of Table 3). Tests of the ACE submodels also exhibit a similar pattern: the model CE dropping the genetic component a^2 cannot be rejected, while both models AE, excluding the shared environment c^2 , and E, excluding both a^2 and c^2 , represent unacceptable deteriorations in fit. We conclude that assortative mating on educational attainment, even at the high level reflected in the $r = .629$ parental correlation, is unlikely to generate the high estimated values of the shared environment component.

We next investigate another mechanism that may be responsible for the unusual pattern of low heritability and high environmentality that we found. Restricting again the discussion to the case of twins, one might conjecture that educational achievement of one twin directly affects educational achievement of the other twin, a process called “phenotypic interaction” in quantitative genetics. For a trait such as smoking, for example, smoking by one twin may directly encourage the behavior in the other twin, through the force of example or through greater availability of cigarettes, net of other genetic or shared environmental influences. With respect to educational attainment two categories of predictions, in opposite directions, are plausible. On one hand it is possible that further educational attainment by one twin would inspire or challenge the other twin to attain the same level; likewise, parents might be particularly concerned to insure equality of educational outcomes among their twins and might thus invest resources in such a way as to compensate for any difference in educational talents between them (Conley 2008). Such *positive interaction* mechanisms would tend to increase the similarity of the twins. On the other hand it is also possible that greater attainment by one twin would tend to reduce attainment by the other. This may happen because the twins, being the same age, would tend to enter college and tax family resources at the same time. This double burden on family finances would result in a negative effect of one twin’s educational attainment on attainment of the other, representing a *negative* or *competitive* interaction.

Reciprocal effects between siblings are notoriously difficult to estimate with ordinary econometric methods (Conley 2008). However, such estimation is possible using an extension of the ACE twin design (Neale and Cardon 1992; Neale and Maes 2004). The model has been shown earlier (Figure 2). Table 8 shows estimates of the

Table 8: ACE+s Models With Siblings Interactions Fit to the Educational Achievement Data – Twins Only (N = 524 Pairs)

Model	Parameter Estimates				Fit Statistics				
	s	a^2	c^2	e^2	χ^2	df	p	AIC	RMSEA
1. ACE+s	.022	.160	.447	.393	1.684	2	.431	-2.316	.026
2. CE+s	.033	— ^a	.543	.457	3.917	3	.271	-2.083	.035
3. AE+s	.243	.257	— ^a	.744	1.922	3	.589	-4.078	.013
4. E+s	.325	— ^a	— ^a	1.0	3.917	4	.417	-4.083	.025
5. ACE	— ^a	.145	.484	.366	1.685	3	.640	-4.315	.007

Note: Parameters are s , phenotypic reciprocal effect; a^2 , additive genetic; c^2 , shared environment; e^2 , nonshared environment.

^a Parameter fixed to value zero.

sibling interaction model for the twins data only.¹¹ The new parameter s in Table 8 estimates the strength of reciprocal influences between siblings. Influences of the twins on each other are assumed to be symmetric. Estimation of the model depends in a crucial way on the use of covariance matrices (as opposed to correlations) in estimation of the model, as the empirical implications of phenotypic interactions consist of differences in the phenotypic variances between MZ and DZ twins, which are lost when using correlations (Neale and Maes 2004). Model ACE+s on Line 1, which combines phenotypic interactions with a full ACE decomposition of the phenotypic variance, satisfactorily fits the data ($\chi^2 = 1.684$ with 2 df, $p = .431$). However the estimate of the reciprocal effect s is small (.022). Dropping the s parameter (model ACE on line 5) does not produce a significant deterioration in fit. Dropping the genetic component a^2 (model CE+s) likewise does not significantly reduce the fit; the estimate of s remains small. Interestingly, dropping the shared environment component c^2 (model AE+s), while it does not significantly reduce the fit, results in a considerable increase (to .243) in the estimated value of s . This pattern suggests that the shared environment c^2 and the reciprocal interaction s in some way represent alternative explanations of the same covariance. Finally, dropping both a^2 and c^2 , while also nonsignificant, results in a further increase (to .325) in the estimated s . Overall, we conclude that including direct phenotypic interaction does not improve the fit of the model.

5 Discussion and Conclusion

Estimating the classical ACE model for educational attainment using data on six groups of twins and nontwin siblings yielded a preferred model that includes moderate heritability ($a^2 = .228$) and substantial shared environmentality, greater among twins ($c^2_{\text{twi}} = .411$) than among nontwin siblings ($c^2_{\text{sib}} = .298$). We interpreted these results as reflecting a rather modest role of native endowment (indicated by the low heritability) and a considerable degree of inequality of opportunity across families (corresponding to the high shared environment estimates). These represent atypical results for a measure of educational achievement that may be assumed to be related to cognitive ability. IQ scores, and presumably scores on other tests strongly related to intelligence such as the SAT and the ACT, typically exhibit much greater heritability, and especially a much lower effect of the shared environment when estimated from young adult twins (e.g., Boomsma, Busjahn and Peltonen 2000). The

¹¹We restrict the model to twins, as a more general model for siblings differing in age is much harder to specify, as direct phenotypic influences of nontwin siblings on each other are less likely to be symmetric.

23 percent estimate for heritability we obtain may not appear too far apart from the admittedly crude estimate of “two fifths” for heritability of educational achievement that Jencks and Tach (2006) propose based on their review of a broad sample of studies, most of them not twin studies. However the persistent role of the shared environment in educational attainment for these young adults, corresponding to 41 percent of the variance in attainment (for twins) and 30 percent (for nontwin siblings), is especially surprising.

The empirical basis for the stronger estimated effect of the shared environment of twins, compared to that of nontwin siblings, is the high value of the DZ correlation (.559) as compared with the correlation for full siblings (.446). According to the classic ACE model correlations for DZ twins and full siblings should be similar, since the two types of siblings have the same degree of biological relatedness. One possibility we explored is that the correlation among twins is inflated due to direct reciprocal influences, called phenotypic interaction, between the twins. While no significant phenotypic interaction effect was found when the full ACE model is included, a significant and strong interaction effect emerged with the AE model (with the shared environment component C excluded). This suggests that phenotypic interaction tends to mimic the effect of the shared environment in the twins data, so that the effects of shared environment and phenotypic interaction are hard to disentangle. This is consistent with much research that has failed to find significant phenotypic interaction (Neale and Maes 2004). This finding is also theoretically consistent, as plausible mechanisms of reciprocal influence of twins on each other would include both positive influences related to emulation and modeling and negative influence related to competition for resources, with the net influence summing up to zero.

We also investigated the potential role of assortative mating in simultaneously deflating the estimate of heritability and inflating the effect of the shared environment. This hypothesis has face plausibility given that assortative mating, measured as the correlation of parental educational attainment, is quite high in the data ($\hat{r} = .629$). However, allowing for assortative mating in the model for twins did not lead to substantially different estimates of the roles of heredity and the shared environment.

The relatively low heritability and high shared environmentality that we found are in sharp contrast with findings of other studies the genetics of educational attainment. For example Behrman and Taubman (1989), using U.S. data with a variety of kinship correlations of educational attainment for twins born 1917–1927 and their parents and offspring, estimate the proportion of variance in schooling “arising from genetic variations” at 81 percent. These authors are surprised to find higher heritability in this data set that includes the offspring of the twins (average age 27 at the time of the survey in 1981–82) than they earlier did based on data for the twins alone (about 40 percent, see Taubman 1976), who were all veterans and thus eligible for the GI Bill, which would have presumably enhanced their educational opportunities, increasing heritability. Likewise, an Australian study of educational attainment using a large sample of twins estimates heritability at 49 percent and shared environmentality at 22 percent, with heritability increasing to 65 percent when an adjustment is made for assortative mating among parents of the twins (Miller, Mulvey and Martin 2001; see also Martin 1978). The high degree of shared environmental variance that we find in this young adult cohort is thus puzzling, in comparison both with the older U.S. cohorts studied by Behrman and Taubman (1989) and the findings from the different, but culturally similar Australian social setting (Miller, Mulvey and Martin 2001; see also Baker et al. 1996).

The system of tertiary education in the U.S. is highly diverse, consisting of a whole range of institutions, some public and some private, that differ widely in

the cost of education for students and their families. A complex assortment of scholarships and financial aid programs further complicates the picture. Our finding of a large shared environment effect suggests that, overall, the pursuit of a university education in the contemporary U.S. is more likely to be constrained by availability of resources on the part of students and their families. The pattern of a persistent “mechanical” effects of family resources on the chances of entering, and graduating from, college. To the extent that our results can be compared to those of Behrman and Taubman (1989), which are based on a somewhat different methodology, they suggest that there has been a considerable worsening in inequality of educational opportunity in the last few decades.¹² We do not have a theory at present why there would be such a trend of increasing inequality of educational opportunity in the U.S.

The view that high inequality of opportunity is related to increasing financial cost of education has implications for comparative research. If one were to systematically compare the results for U.S. data presented in this paper with comparable quantitative genetic data for other advanced societies that have essentially free or at least less expensive systems of tertiary education, one should find that the shared environment component of educational attainment is larger in the U.S. (and heritability correspondingly smaller) than in these other countries. In such comparisons the size of the shared environment constitutes a measure of inequality in schooling opportunity capturing the influence of social privilege on attainment while conversely heritability measures the role of endowment (native potential) in attainment (Behrman and Taubman 1989). Heath et al. (1985) provide an illuminating example of the useful macrosociological implications of the quantitative genetic model for such comparisons. Based on large samples of Norwegian twins they find that the heritability of educational attainment has increased (to 67–74 percent) for younger cohorts of males (born 1950 or later) compare to older cohorts (born before 1940), and the role of the shared environment has correspondingly decreased (to 8–10 percent). They see this trend as reflecting democratic reforms in the Norwegian educational system after World War II. However heritability of schooling has not increased for females over the same period, from which they conclude that equality of opportunity for females has not improved to the same extent as for males, so the role of family privilege in educational attainment remains high for females (see also Baker et al. 1996).

One general implication of the quantitative genetic analysis presented in this paper is that, as recently and forcefully argued by Jencks and Tach (2006), the quantitative genetic decomposition of the variance in educational attainment provides a means to assess the degree of inequality of opportunity in schooling (and socioeconomic success in general) in a society. Such a measure had so far proved elusive in the social stratification literature, as the simple intergenerational association of status, measured as a parent-offspring correlation or as diagonal cells disproportion in a mobility table, is unsuitable as a measure of opportunity (Eckland 1967; see also Guo and Stearns 2002; Nielsen 2006). Inequality of schooling opportunity measured as the proportion of shared environment variance and, conversely, the role of endowment (merit) measured as heritability can then be correlated with other societal characteristics such as the dominant political ideology, the nature of the welfare system, or degree of modernization to test longstanding comparative theories of social stratification (Hout and DiPrete 2006).

A major puzzle implicit in our results, worth exploring in future research, is the apparent discrepancy between the pattern of low heritability / high shared

¹²Behrman and Taubman’s (1989) offspring of twins were 27 years old on average in 1981–82, which would make them roughly 30 years older than the AddHealth cohort.

environmentality that we find for educational attainment and the opposite pattern that is commonly found for IQ and other cognitively-loaded measures. Based on a large Dutch sample of late adolescent twins, for example, Boomsma, Busjahn and Peltonen (2002) find heritabilities greater than 80 percent and no shared environmentality. Using the same AddHealth siblings in the first two waves, Nielsen (2006) likewise finds heritabilities of 54 percent, 67 percent and 60 percent (and correspondingly low shared environmentalities) for verbal IQ, high school GPA, and college plans (a composite measure of college aspiration and expectation). To the extent that a measure of educational achievement reflects cognitive ability, as has been well documented for such measures as scores on national examinations (e.g., Deary et al. 2007), one would expect heritability to be relatively high and shared environmentality to be relatively low, as most often found for IQ. The high heritability pattern makes sense as scores on standardized tests cannot usually be “mechanically” manipulated in resources-dependent ways by individuals and their families, and softer, cultural influences of the family tend to fade by late adolescence (Nielsen 2008; Rowe 1994). Comparison of measures of educational achievement that differ in their cognitive loadings — such as GPA, scores on standardized tests such as the SAT, or even the selectivity of the college attended — should reveal systematic variation in the roles of heredity versus the shared environment, with the more cognitively-loaded outcomes showing greater heritability and a reduced role of the shared environment.

Highest degree earned or years of education completed constitute rather undifferentiated measures of educational attainment that do not distinguish degrees according to the selectivity, prestige or “quality” (however defined) of the institution where it is earned, so that a degree from a nonselective college and from an elite university (and the years spent earning it) are rated the same with respect to attainment. One could speculate that if one had a measure of attainment incorporating the “quality” of the degree earned, and to the extent that entering and graduating from a higher quality institution is more cognitively demanding, such a measure would behave from a quantitative genetic viewpoint more like a cognitive measure. The very strong effect of standardized test scores (such as the SAT or ACT) on admission to more selective institutions is consistent with that scenario (Briggs 2001; Buchmann, Condrón and Roscigno 2010). One would predict that the quantitative genetic structure of such a quality-weighted measure of educational attainment would tend to resemble that of IQ, with a stronger effect of genes (heritability) and a weaker impact of the shared environment.

Measuring educational attainment as highest degree earned, as we did in this paper, provides an overall, summary measure of an individual’s educational career that combines together several discrete transitions (Mare 1981). Using similar types of quantitative genetic models (adapted for the dichotomous outcomes) it is possible to estimate the respective roles of genes and of the shared environment for each transition, such as entering college, graduating, and pursuing a post-graduate degree, to form a more precise picture of where exactly in the educational career genes and the family environment matter most. The hope for such nuanced insights is indeed the principal motivation for Mare’s (1981) focus on educational transitions. The added benefit of the quantitative genetic approach is that it provides clear measures of the respective roles of individual merit (heritability) versus family privilege (shared environmentality). Thus estimating the quantitative genetic model for transitions separately might help identifying “opportunity bottlenecks”, transitions with low heritability / high shared environmentality, that could become targets for further research and perhaps policy attention.

The analyses presented in this paper use data on nontwin siblings in addition to the data on MZ and DZ twins that constitute the core of the classical twin design. We found that the quantitative models for twins and for nontwin siblings differ signifi-

cantly, particularly with respect to the greater impact of the shared environment of twins. While we have mentioned some of the likely mechanisms, such as the greater likelihood of resources competition due to contemporaneity, that differentiate twins from other types of siblings, further research is needed to refine our understanding of the relationship between twins and nontwin siblings that underlie this extended design. The need for a better understanding would seem to be particularly acute with respect to various measures of educational achievement, which represent the outcomes of processes deployed over time, involving the accumulation of interactive choices by the siblings as well as their families.

The main findings of moderate heritability of educational attainment and a larger role of the shared environment, while obtained with very standard techniques of quantitative genetic analysis, are relatively unexpected as most human traits tend to conform to Turkheimer's (2000) "second law" in exhibiting stronger heritability than shared environmentality. In the present context the empirical pattern of moderate genetic influences and strong shared environmentality indicates that substantial inequality of opportunity for educational attainment persists (or, perhaps, has newly developed) in American society at the turn of the twenty first century. While we had theoretical reasons to expect that schooling measured as highest degree earned might show a larger, resources-related effect of the shared environment than more cognitive measures of educational success, such as grades or test scores, we did not expect, nor did we deliberately construct, the lopsided shared environmental determination of educational attainment that we found.

The empirical pattern that we extracted from the data using standard analytical methodology, if it holds to further scrutiny, indicates a strong persistence in inequality of schooling opportunity in the contemporary U.S. This pattern implies at least two kinds of theoretical challenges. One challenge is in understanding why inequality of schooling opportunity appears to be so high in such a recent period of American history, when so much attention and resources have ostensibly been devoted to equalizing opportunities for members of different gender and ethnic groups and social classes. The other challenge is in understanding how educational attainment, as we have measured it as highest degree earned, relates to the quantitative genetics of other traits, such as cognitive ability and personality factors, that appear to depend on genes to a larger extent. There is a need to reconcile findings of a considerable role of genes (high heritability) in such stratification-related traits as IQ and some personality factors such as conscientiousness, and the larger role of the shared environment we found for educational attainment, given that these cognitive and personality traits more probably contribute substantially to educational attainment. There is therefore a need to unravel the genetic and environmental underpinnings of the relationships between attainment-related traits (see e.g. Nielsen 2006).

Finally, we hope that our study has contributed to illustrating how the quantitative genetic decomposition of the variance in educational attainment more generally provides ways to measure central concepts of social stratification, such as inequality of opportunity (measured as the proportion of the variance due to the shared environment) versus the degree of meritocracy associated with the role of genes (measured as the proportion of variance due to genes, or heritability) which have eluded purely phenotypically-oriented research traditions.

References

- Baker, Laura A., Susan A. Treloar, Chandra A. Reynolds, Andrew C. Heath, and Nicholas G. Martin. 1996. "Genetics of Educational Attainment in Australian Twins: Sex Differences and Secular Changes." *Behavior Genetics* 26: 2: 89–102.
- Behrman, Jere R. and Paul Taubman. 1989. "Is Schooling Mostly in the Genes?"

- Nature-nurture Decomposition Using Data on Relatives.” *Journal of Political Economy* 97: 6: 1425–1446.
- Beller, Emily and Michael Hout. 2006. “Welfare states and social mobility: How educational and social policy may affect cross-national differences in the association between occupational origins and destinations.” *Research in Social Stratification and Mobility* 24: 353–365.
- Blau, Peter M., and Otis Dudley Duncan. 1967. *The American Occupational Structure*. New York: Wiley.
- Blossfeld, Hans-Peter and A. Timm, eds. 2003. *Who Marries Whom? Educational Systems as Marriage Markets in Modern Societies*. Dordrecht, Netherlands: Kluwer.
- Bollen, Kenneth A. 1989. *Structural Equations with Latent Variables*. New York: Wiley.
- Boomsma, Dorret, Andreas Busjahn and Leena Peltonen. 2002. “Classical Twin Studies and Beyond.” *Nature Reviews Genetics* 3: 872–882.
- Bouchard, Thomas J., Jr. and Matt McGue. 2003. “Genetic and Environmental Influences on Human Psychological Differences.” *Journal of Neurobiology* 54: 1 (Jan): 4–45.
- Boudon, Raymond. 1974. *Education, Opportunity, and Social Inequality*. New York: John Wiley.
- Breen, Richard, and Jan Jonsson. 2005. “Inequality of Opportunity in Comparative Perspective: Recent Research on Educational Attainment and Social Mobility.” *Annual Review of Sociology* 31: 223–243.
- Briggs, Derek C. 2001. “The Effect of Admissions Test Preparation: Evidence from NELS:88” *Chance* 14: 1: 10–22.
- Buchmann, Claudia, Dennis J. Condrón and Vincent J. Roscigno. 2010. “Shadow Education, American Style: Test Preparation, the SAT and College Enrollment.” *Social Forces* 89: 2(December): 435–461.
- Buchmann, Claudia and Thomas A. DiPrete. 2006. “The Growing Female Advantage in College Completion: The Role of Parental Resources and Academic Achievement.” *American Sociological Review* 71: 515–541.
- Carey, Gregory. 2003. *Human Genetics for the Social Sciences*. Thousand Oaks, CA: Sage.
- Cherny, S. S. 2008. “Variance Components and Related Methods for Mapping Quantitative Trait Loci.” *Sociological Methods and Research* 17: 2 (Nov): 227–250.
- Chipuer, Heather M., Michael J. Rovine, and Robert Plomin. 1990. “LISREL Modeling: Genetic and Environment Influences on IQ Revisited.” *Intelligence* 14 (1): 11–29.
- Conley, Dalton. 2008. “What Do Low (or High) Sibling Correlations Tell Us About Social Ascription?” Pp. 596–609 in *Social Stratification* (3d edition), edited by David B. Grusky. Boulder, CO: Westview Press.
- Deary, Ian J., Steve Strand, Pauline Smith, and Cres Fernandes. 2007. “Intelligence and Educational Achievement”. *Intelligence* 35: 1(January-February): 13–21.
- Duncan, Otis Dudley, David L. Featherman and Beverly Duncan. 1972. *Socioeconomic Background and Achievement*. Seminar Press.

- Duncan, Otis Dudley and Robert W. Hodge. 1963. "Education and Occupational Mobility: A Regression Analysis." *American Journal of Sociology* 68: 6: 629–644.
- Eaves, Lindon J., Heath, Andrew C., and Martin, Nicholas G. 1984. "A Note on the Generalized Effects of Assortative Mating." *Behavior Genetics* 14: 371–376.
- Eaves, Lindon, Andrew Heath, Nicholas Martin, Hermine M. Maes, Michael C. Neale, Kenneth Kendler, Katherine Kirk and Linda Corey. 1999. "Comparing the Biological and Cultural Inheritance of Personality and Social Attitudes in the Virginia 30,000 Study of Twins and Their Relatives." *Twin Research* 2: 62–80.
- Eckland, Bruce K. 1967. "Genetics and Sociology: A Reconsideration." *American Sociological Review* 32: 173–194.
- Eckland, Bruce K. 1979. "Genetic Variance in the SES-IQ Correlation." *Sociology of Education* 52: 3: 191–196.
- Falconer, D. S. 1989. *Introduction to Quantitative Genetics*. 3d edition. New York: John Wiley & Sons.
- Fisher, Ronald A. 1918. "The Correlation between Relatives on the Supposition of Mendelian Inheritance." *Transactions of the Royal Society of Edinburgh* 52, Part 2: 399–433.
- Freese, Jeremy. 2008. "Genetics and the Social Science Explanation of Individual Outcomes." *American Journal of Sociology* 114 Supplement: S1–S35.
- Freese, Jeremy and Sara Shostak. 2009. "Genetics and Social Inquiry." *Annual Review of Sociology* 35: 107–128.
- Frey, Meredith C. and Douglas K. Detterman. 2004. "Scholastic Assessment or g? The Relationship Between the Scholastic Assessment Test and General Cognitive Ability." *Psychological Science* 15: 6 (): 373–378.
- Galton, Francis. 1869. *Hereditary Genius*. London, England: Macmillan.
- Ganzeboom, Harry B. G., Donald J. Treiman, and Wout C. Ultee. 1991. "Comparative Intergenerational Stratification Research: Three Generations and Beyond." *Annual Review of Sociology* 17: 277–302.
- Guo, Guang, and Elizabeth Stearns. 2002. "The Social Influences on the Realization of Genetic Potential for Intellectual Development." *Social Forces* 80: 881–910."
- Harris, Kathleen Mullan. 2009. *The National Longitudinal Study of Adolescent Health (Add Health), Waves I & II, 1994–1996; Wave III, 2001–2002; Wave IV, 2007–2009* [machine-readable data file and documentation]. Chapel Hill, NC: Carolina Population Center, University of North Carolina at Chapel Hill.
- Harris, Kathleen Mullan, Carolyn Tucker Halpern, Andrew Smolen and Brett C. Haberstick. 2006. "The National Longitudinal Study of Adolescent Health (Add Health) Twin Data." *Twin Research and Human Genetics* 9: 6: 988–997.
- Heath, A. C., K. Berg, L. J. Eaves, M. H. Solaas, L. A. Corey, J. Sundet, P. Magnus and W. E. Nance. 1985. "Education Policy and the Heritability of Educational Attainment." *Nature* 314 (6013): 734–736.
- Hout, Michael and Thomas A. DiPrete. 2006. "What Have We Learned: RC28's Contributions to Knowledge about Social Stratification." *Research in Social Stratification and Mobility* 24: 1–20.

- Jencks, Christopher, M. Smith, H. Acland, M. J. Bane, D. Cohen, H. Gintis, B. Heyns, and S. Michelson. 1972. *Inequality: A Reassessment of the Effect of Family and Schooling in America*. New York: Basic Books.
- Jencks, Christopher and Laura Tach. 2006. "Would Equal Opportunity Mean More Mobility?" Pp. 23–58 in *Mobility and Inequality: Frontiers of Research from Sociology and Economics*, edited by Stephen Morgan, David Grusky and Gary Fields. Stanford, CA: Stanford University Press.
- Jensen, Arthur R. 1998. *The g Factor: The Science of Mental Ability*. Westport, CT: Praeger.
- Jinks, J. L. and D. W. Fulker. 1970. "Comparison of the Biometrical Genetical, MAVA, and Classical Approaches to the Analysis of Human Behavior." *Psychological Bulletin* 73: 5 (May): 311–349.
- Koenig, Katherine A., Meredith C. Frey and Douglas K. Detterman. 2008. "ACT and general cognitive ability." *Intelligence* 36: 2 (March-April): 153–160.
- Kohler, Hans-Peter, Jere R. Behrman and Jason Schnittker. 2011. "Social Science Methods for Twins Data: Integrating Causality, Endowments, and Heritability." *Biodemography and Social Biology* 57: 1: 88–141.
- Lichtenstein, Paul, Nancy L. Pedersen and G. E. McClearn. 1992. "The Origins of Individual Differences in Occupational Status and Educational Level: A Study of Twins Reared Apart and Together." *Acta Sociologica* 35: 1: 13–31.
- Mare, Robert D. 1980. "Social Background and School Continuation Decisions." *Journal of the American Statistical Association* 75: 370 (June): 295–305.
- Mare, Robert D. 1981. "Change and Stability in Educational Stratification." *American Sociological Review* 46: 1 (February): 72–87.
- Mare, Robert D. 1991. "Five Decades of Educational Assortative Mating." *American Sociological Review* 56 (1): 15–32.
- Martin, Nicholas G. 1978. "Genetics of Sexual and Social Attitudes in Twins." Pp. 13–23 in *Twin Research: Proceedings of the Second International Congress on Twin Studies, August 29 – September 1, 1977, Washington, D.C., Part A: Psychology and Methodology*, edited by Walter E. Nance. New York: Alan R. Liss.
- Martin, N. G. and L. J. Eaves. 1977. "The Genetical Analysis of Covariance Structure." *Heredity* 38: 79–95.
- Mayer, Susan E. 1997. *What Money Can't Buy: Family Income and Children's Life Chances*. Cambridge, MA: Harvard University Press.
- Medland, Sarah E. and Peter K. Hatemi. 2009. "Political Science, Biometric Theory, and Twin Studies: A Methodological Introduction." *Political Analysis* 17: 2 (Spring): 191–214.
- Miller, Paul, Charles Mulvey and Nicholas G. Martin. 2001. "Genetic and Environmental Contributions to Educational Attainment in Australia." *Economics of Education Review* 20: 3 (June): 211–224.
- Neale, Michael C., Steven M. Boker, Gary Xie, and Hermine H. Maes. 2003. *Mx: Statistical Modeling*. 6th ed. Richmond: Virginia Commonwealth University, Department of Psychiatry.

- Neale, Michael C. and Lon R Cardon. 1992. *Methodology for Genetic Studies of Twins and Families*. NATO ASI Series. Dordrecht, Netherlands: Kluwer Academic Press.
- Neale, Michael C. and Hermine M. Maes. 2004. *Methodology for Genetic Studies of Twins and Families*. Draft 2nd edition. Dordrecht, Netherlands: Kluwer Academic Press.
- Neale, Michael C. and Michael B. Miller. 1997. "The Use of Likelihood-Based Confidence Intervals in Genetic Models." *Behavior Genetics* 27: 2: 113–120.
- Nielsen, François. 2006. "Achievement and Ascription in Educational Attainment: Genetic and Environmental Influences on Adolescent Schooling." In *The Linking of Sociology and Biology* [special section], edited by Guang Guo. *Social Forces* 85: 1 (September): 193–216.
- Nielsen, François. 2008. "The Nature of Social Reproduction: Two Paradigms of Social Mobility." Symposium / Social Sciences and Natural Sciences – What Connection?, edited by Mario Lucchini and Maurizio Pisati. *Sociologica* Nr 3 / 2008 (November): Pp. 1–35. [Online at <http://www.sociologica.mulino.it/>.]
- Petrill, Stephen A. and Bessie Wilkerson. 2000. "Intelligence and Achievement: A Behavioral Genetic Perspective." *Educational Psychology Review* 12: 2 (June): 185–199.
- Rosenfeld, Michael. 2008. "Racial, Educational and Religious Endogamy in the United States: A Comparative Historical Perspective." *Social Forces* 87: 1 (Sep): 1–31.
- Rowe, David C. 1994. *The Limits of Family Influence: Genes, Experience, and Behavior*. Guilford.
- Rowe, David C. and Kristen C. Jacobson. 1998. National Longitudinal Study of Adolescent Health: Pairs Code Book. Accessed at: <http://www.cpc.unc.edu/addhealth/codebk1.html/pairs.zip>
- Scarr-Salapatek, Sandra. 1971. "Race, Social Class, and IQ." *Science* 174: 1285–95.
- Smits, Jeroen, Wout Ultee and Jan Lammers. 1998. "Educational Homogamy in 65 Countries: An Explanation of Differences in Openness Using Country-Level Explanatory Variables." *American Sociological Review* 63: 2 (April): 264–285.
- Solon, Gary. 2008. "Intergenerational Income Mobility." Pp. 481–485 in *Social Stratification: Class, Race, and Gender in Sociological Perspective*, edited by David B. Grusky. Boulder, CO: Westview Press.
- Tambs, Kristian, Jon Martin Sundet, Per Magnus and Kare Berg. 1989. "Genetic and Environmental Contributions to the Covariance Between Occupational Status, Educational Attainment, and IQ: A Study of Twins." *Behavior Genetics* 19: 209–222.
- Taubman, Paul. 1976. "The Determinants of Earnings: Genetics, Family and Other Environments: A Study of White Male Twins." *American Economic Review* 66: 5: 858–870.
- Turkheimer, Eric. 2000. "Three Laws of Behavior Genetics and What They Mean." *Current Directions in Psychological Science* 9: 5 (Oct): 160–164.
- Visscher, Peter M., William G. Hill and Naomi R. Wray. 2008. "Heritability in the Genomics Era: Concepts and Misconceptions." *Nature Reviews Genetics* 9: a (April): 255–266.

- Vogler, George P. and David W. Fulker. 1983. "Familial Resemblance for Educational Attainment." *Behavior Genetics* 13: 4: 341-354.
- Wright, Sewall. 1921a. "Correlation and Causation." *Journal of Agricultural Research* 20: 557-585.
- Wright, Sewall. 1921b. "Systems of Mating. I. The Biometric Relations between Parents and Offspring." *Genetics* 6: 111-123.
- Wright, Sewall. 1934. "The Method of Path Coefficients." *Annals of Mathematical Statistics* 5: 161-215.