SOCIAL
FORSKNINGS
INSTITUTTET

Mads Meier Jæger

05:2006 WORKING PAPER

DOES SIBSHIP SIZE AFFECT EDUCATIONAL ATTAINMENT?

A RE-EVALUATION OF THE RESSOURCE DILUTION HYPOTHESIS USING INSTRUMENTAL VARIABLES

RESEARCH DEPARTMENT OF SOCIAL POLICY AND WELFARE

DOES SIBSHIP SIZE AFFECT EDUCATIONAL ATTAINMENT?

A RE-EVALUATION OF THE RESOURCE DILUTION HYPOTHESIS USING INSTRUMENTAL VARIABLES

Mads Meier Jæger

Social policy and welfare
Working Paper 05:2006

Cover page

Does Sibship Size Affect Educational Attainment?

A Re-evaluation of the Resource Dilution Hypothesis Using Instrumental Variables

Mads Meier Jæger*

The Danish National Institute of Social Research

Abstract

This paper implements a test of the Resource Dilution Hypothesis (RDH) stating that sibship size has a negative causal effect on educational attainment. Most existing studies using conventional methods support the RDH. This paper implements an Instrumental Variable (IV) approach to testing the claim of a negative causal relationship between sibship size and educational attainment. Analyzing data from the Wisconsin Longitudinal Study, the empirical analysis demonstrates, first, that conventional OLS regression estimates sibship size to have a negative effect on educational attainment equal to about one-tenth of a year of schooling per sibling. Second, when applying the IV method to account for potential endogeneity, the negative effect of sibship size increases substantially to about one-third of a year of schooling per sibling.

Word count: 8,079 (including in-text references and endnotes).

* The Danish National Institute of Social Research, Herluf Trolles Gade 11, DK-1052 Copenhagen K. E-mail: mads@sfi.dk. The research reported in this paper was sponsored by the Danish Social Science Research Council.

Running head: Sibship size and educational attainment

One of the most robust findings in the literature on family background and educational outcomes is that sibship size is negatively associated with children's intellectual and educational outcomes. In fact, among the sibship characteristics typically studied in the literature: sibship size, birth order, birth spacing, and sibship sex composition, only the negative relationship between sibship size and intellectual and educational outcomes has been consistently reproduced over time and across a range of Western, industrialized countries (for reviews of the literature see Cicirelli 1978; Ernst and Angst 1983; Heer 1985; Steelman 1985; Haveman and Wolfe 1995; Steelman et al. 2002). Kuo and Hauser (1997:73) describe the negative relationship between sibship size and children's outcomes as "inarguable", while Steelman et al. (2002:249) in their review state that the relationship " … typically persists regardless of educational outcome, be it performance on standardized exams, grades in school, educational expectations and aspirations, or educational attainment".

The leading theoretical perspective developed to account for this negative relationship is the *Resource Dilution Hypothesis*. This hypothesis states that parental resources: money, time, care etc., are limited, and when the size of the sibship grows the amount of resources available to each child in the family becomes increasingly diluted (Blake 1981, 1985, 1989; Downey 1995, 2001). As a consequence, children from large families exhibit poorer intellectual ability and lower levels of educational attainment compared to children from smaller families. However, in recent years critical voices have been raised against the Resource Dilution Hypothesis. In particular, the claim that the negative relationship between sibship size and intellectual and educational attainment constitutes a true *causal* relationship has been questioned (see Ernst and Angst 1983; Heer 1985; Guo and VanWey 1999a; Rodgers 2001; Conley and Glauber 2005). Critics argue that the negative effect of sibship size found in most empirical studies using cross-sectional data could be spurious, meaning that rather than capturing a true causal effect of sibship size on educational attainment the variable measuring sibship size actually captures unmeasured socioeconomic or familial traits. Proponents of the Resource Dilution Hypothesis have acknowledged the possibility that the negative effect of sibship size on intellectual and educational outcomes found in most empirical studies may be overstated (Downey 2001:502; Steelman et al. 2002:253-256), although only very few studies have so far attempted to test this possibility directly.

The objective of this paper is to test if sibship size, as claimed in the Resource Dilution Hypothesis, has an independent and causal effect on children's educational attainment. Using data from the Wisconsin Longitudinal Study, the paper utilizes a statistical technique developed specifically to test causal hypotheses: Instrumental Variable (IV) estimation. The basic idea in IV estimation is that if, as suggested in recent critical studies, sibship size could be endogenous to children's educational attainment (i.e. the estimated effect of sibship size is biased due to sibship size being correlated with unobserved socioeconomic variables), then this endogeneity bias could be neutralized if one could find some "instrumental" variable which determines sibship size but *not* children's educational attainment. By utilizing the variation in sibship size induced by an instrumental variable strictly exogenous to children's educational attainment, one would be able to obtain an unbiased estimate of the true causal effect of sibship size on children's educational attainment. The conditions and assumptions under which IV methods work are described later in the paper.

In recent years, several attempts have been made to implement tests of the hypothesized causal relationship between sibship size and educational outcomes. Guo and VanWey (1999a) use panel data and a fixed effect model to correct for the influence of time-invariant unobserved variables, while Baydar et al. (1997) proposes a model combining level and time-varying explanatory variables. Both studies attempt to deal with unobserved family characteristics, and, as I discuss below, both studies to some extent provide a more plausible estimate of the causal effect of sibship size on educational outcomes than do conventional approaches. However, both studies are also limited in that they do not carry out any formalized tests as to ascertain that their alternative estimates (compared to those obtained using cross-sectional data and methods) are unbiased. This leaves open the question of whether or not they succeed in their objective. In addition, two recent studies have used IV methods to investigate the causal relationship between sibship size and educational attainment (Black et al. 2005; Conley and Glauber 2005). IV methods are attractive since they provide a formalized framework within which to test causal hypotheses. In this paper I adopt the IV framework but improve upon the existing IV studies in two key areas by (1) introducing instrumental variables that have a better theoretical justification compared to those used in previous studies, and (2) carrying out formalized tests of the validity of the instrumental variables.

With respect to the first improvement, previous IV studies typically use the sex composition of the sibship (or sometimes the occurrence of twin births) as an instrumental variable providing an exogenous source of variation in sibship size. This exogenous variation is then used to identify the causal effect of sibship size on educational attainment (Black et al. 2005; Conley and Glauber 2005). However, in order for sex composition to be a theoretically meaningful instrumental variable, most studies have to assume that parents plan fertility based on the "quality" (in terms of education, earnings potential, etc.) of children already born. Typically, boys are hypothesized to be more "productive" than girls (e.g. Becker and Tomes 1986). This behavioral assumption rests on theoretical grounds and is not (and cannot be) tested in empirical applications. In this paper I take a different approach and use as instrumental variables four indicators capturing parents' genetic or cultural propensity to have large sibships. These four indicators are (1) the size of the sibship from which both parents originate, and (2) the age of both parents at the time of the birth of their first child. Demographic research shows that both types of indicators constitute useful proxies for parents' inherited reproductive capability and "environmentally" (e.g. socially, culturally, and religiously) induced preferences for family size (e.g. Michael and Tuma 1985; Axinn et al. 1994; Murphy and Knudsen 2002). Since individuals who form marriages and eventually have children are initially unaware of the genetic reproductive capabilities of their partners (although they will probably be partly aware of partners' cultural preferences for certain family types), the instruments used in this paper comprise a "natural" experiment in the sense that parents are randomly "assigned" into families with partners with different reproductive capabilities. Consequently, my instrumental variables need not assume that parents plan fertility based on a theory of the expected productivity of their offspring. Rather, the perspective advocated here only presupposes, first, that a variation in individuals' reproductive capability and family size preferences exists which affects how many children they have, and, second, that individuals are generally unaware of their partners' genetic reproductive capabilities when they form a couple. These assumptions, I believe, are far less restrictive than those imposed in previous studies. As I demonstrate in the empirical analysis, the instrumental variables proposed are indeed valid.

With respect to the second improvement over previous IV studies, in this study I use multiple rather than a single instrumental variable. Having multiple instrumental variables provides the possibility of evaluating the empirical validity of the IV procedure. Since a key property of instrumental variables, in addition to predicting the endogenous variable under study, is that they must be exogenous with respect to the outcome variable, the possibility of carrying out formalized statistical tests of this assumption is crucial. Previous IV studies using single rather than multiple instruments (see Black et al. 2005; Conley and Glauber 2005) are not able to carry out such tests, which means that the validity of one of the basic motivations for deploying IV methods to solve the inferential problem regarding the causal relationship between sibship size and educational attainment remains unresolved. Using data from the Wisconsin Longitudinal Study (WLS), I find that standard OLS estimation of the effect of sibship size on the number of years of schooling completed by the children of the WLS respondents produces a significant negative effect. The size of this negative effect (about one-tenth of a year of schooling per sibling) is comparable to that found in previous studies. When applying the IV method I find that once potential endogeneity is accounted for, the negative causal effect of sibship size on educational attainment is in fact much stronger (about one-third of a year of schooling per sibling), thereby providing strong empirical support for the Resource Dilution Hypothesis. I attribute these considerable differences in the estimated effects of sibship size on educational attainment to selection mechanisms; i.e. that for families which have many children large sibships may not necessarily be as “hurtful” to children’s educational careers compared to “average” families (for which the IV estimate applies).

The paper proceeds as follows. First, in the next section I present the Resource Dilution Hypothesis. This section also discusses findings from previous studies. Second, I present the Wisconsin Longitudinal Study and the variables and measures used in my analysis. Third, I outline the empirical framework used in the paper. Forth, following this presentation I carry out the empirical analysis of the causal effect of sibship size on children’s educational attainment. In the final section of the paper I conclude on the empirical analysis and discuss some avenues for future research.

The Resource Dilution Hypothesis

Several theoretical approaches have been developed to explain the negative relationship between family size and intellectual and educational outcomes. In this section I present the Resource Dilution Hypothesis which in recent years has become the leading theoretical explanation of the relationship between sibship size and intellectual and educational developments (Blake 1981, 1985, 1989; Downey 1995, 2001; Steelman et al. 2002).[1] Furthermore, I also provide a review of the existing empirical literature analyzing the relationship between sibship size and educational outcomes.

The Resource Dilution Hypothesis

The Resource Dilution model starts from the observation that parental resources: financial, physical and emotional, are intrinsically limited. Parents may provide three types of resources to children: (1) *physical settings* (a home, food, cultural objects like books and works of art); (2) *treatments* (personal attention, teaching); and (3) *opportunities* (chances to interact with the surroundings and the outside world) (Blake 1981; Downey 2001:498). When the size of the family grows, the amount of all types of resources available to each child in the family becomes increasingly diluted, and, *ceteris paribus*, children from large families tend to display lower intellectual capacity and poorer educational performance compared to children from smaller families (Blake 1981, 1989; Downey 1995, 2001). As a consequence, the Resource Dilution Hypothesis predicts that sibship size has an independent and negative causal effect on children’s educational outcomes over and above the effects of other sibship, familial, socioeconomic, and demographic characteristics. Economic theory has developed a similar argument under the slogan of child “quantity” versus “quality” (e.g. Becker and Tomes 1986).

The Resource Dilution Hypothesis then provides a straight-forward hypothesis relating to the *total size* of the sibship. However, three additional sibship characteristics may also play a role with respect to how parents prioritize resources and how many resources children receive (Steelman 1985). First, birth order may be important as early-born children, until the birth of additional siblings, receive proportionally more parental resources, which in turn may provide a comparative advantage later in life (Marjoribanks and Walberg 1975; Galbraith 1982; Ernst and Angst 1983). Second, birth spacing may also play a role since children born in close proximity must share more parental resources compared to siblings born further apart (Rosenzweig 1986; Powell and Steelman 1990, 1993). Finally, the sex composition of the sibship has been hypothesized to play a role since parents may have different preferences for investing in respectively boys and girls (Butcher and Case 1994; Kaestner 1997; Kuo and Hauser 1997; Conley 2000). Although in this paper I focus on sibship size as the key explanatory variable, in the empirical analysis presented below I include indicators of birth order, birth spacing, and sibship sex composition as controls.

TABLE 1 ABOUT HERE

Evidence from Previous Studies

But how important is sibship size for educational attainment?[2] Table 1 summarizes the results of 28 empirical studies which include sibship size as an independent variable explaining educational attainment measured by years of completed schooling. As is evident from the table, practically all studies report a statistically significant negative effect of sibship size on years of completed schooling. The effect of sibship size on educational attainment varies by country, sex, race, cohort, number of control variables included, and research design, but it appears that adding an additional sibling to the family has a detrimental effect on children's educational attainment equivalent to about one-tenth to one-third of a year of schooling. Appendix table 1 reports the findings from an additional 19 empirical studies using other measures of educational outcomes than years of completed schooling (categorical representations of educational attainment, test score results/GPA, etc.), most of which also find that sibship size has a negative effect on educational outcomes. In addition, previous studies based on the Wisconsin Longitudinal Study find a negative effect of sibship size on educational attainment (see Hauser and Sewell 1985; Kuo and Hauser 1997; Plug and Vijverberg 2003). The empirical evidence then seems overwhelming in favor of the Resource Dilution Hypothesis suggesting a negative causal relationship between sibship size and educational attainment.

Alternative Approaches

As mentioned previously, recent studies using more sophisticated methodological approaches have questioned the causal nature of the relationship between sibship size and educational attainment. Most of the studies cited in table 1 (and appendix table 1) use cross-sectional data and conventional statistical methods such as the Ordinary Least Squares (OLS) regression model. Since sibship size may act as a proxy for unobserved family or socioeconomic traits (for example, parents' social, organizational, and intellectual resources), it is possible that OLS provides biased or entirely spurious estimates of the effect of sibship size on educational attainment. If this were the case the Resource Dilution Hypothesis would be compromised.

As an indication that traditional cross-sectional studies may be overestimating the negative effect of sibship size on educational attainment, observe that the studies cited in table 1 which use sibling models to account for observed and unobserved family influences shared by siblings (Olneck and Bills 1979; de Graaf 1986; de Graaf and Huinink 1992; Sieben et al. 2001) tend to report lower (but still significant) negative effects of sibship size on educational attainment compared to traditional cross-sectional studies. One of the first empirical studies to test the causal nature of the relationship between sibship size and (in their case) cognitive ability is Guo and VanWey (1999a) who used longitudinal data and a fixed effect model to take into account time-invariant unobserved (individual and family) heterogeneity. They find that once unobserved variables are controlled, sibship size no longer has a significant effect on children's cognitive test scores. The analysis by Guo and VanWey has been extensively criticized by Downey et al. (1999) and Phillips (1999) (with a response to this criticism in Guo and VanWey (1999b)) especially for the selective empirical sample and measures used. Although some of this criticism concerning the actual (NLSY) sample used is warranted, the principal objections to the methodological approach used by Guo and VanWey are not.[3] However, the small sample size analyzed by Guo and VanWey in combination with the fixed effect model is a source of concern, as there may insufficient information (and, in particular, insufficient births of additional siblings) in the data to identify the causal effect of sibship size on cognitive test scores. As a consequence, the finding of no effect of sibship size reported by Guo and VanWey (1999a) could be the result of the inefficiency of the statistical method used rather than a substantive finding.

Recent studies based on instrumental variable estimation reach conclusions similar to those of Guo and VanWey (1999). Conley and Glauber (2005) provide IV estimates of the likelihood of attending private school and being held back one grade given sibship size (or, more specifically, when going from sibships of size two to size three). They find that IV estimates of the effect of going from two to three children on the likelihood of private school attendance are statistically significant and negative.[4] On the other hand, they find that while the likelihood of being held back one grade increases with sibship size in a standard logit model, then this effect is no longer significant in the

IV model. This result would indicate that the effect of sibship size is spurious in their study, although they do not study the final educational attainment of their respondents. Using Norwegian register data, Black et al. (2005) analyze the impact of sibship size on years of completed schooling. Using OLS regression they find a raw correlation between sibship size and years of completed schooling of -.182; a figure which is comparable to the estimates from previous studies reported in table 1. However, when controlling for demographic and socioeconomic factors, as well as birth order, the effect of sibship size reduces to -.013 (but remains statistically significant). Finally, when deploying the IV method to correct for endogeneity sibship size is no longer significant.

As mentioned above, a problem in the studies by Conley and Glauber (2005) and Black et al. (2005) is that the motivation behind the instrumental variables used assumes that parents plan future fertility based on the sex composition of the sibship or the occurrence of twin births. Usually, there is no way of confirming if these behavioral assumptions hold in practice. Furthermore, when using single rather than multiple instrumental variables it is not possible to test if the instruments are empirically valid in terms of solving the endogeneity problem (this issue is discussed shortly). As a consequence, in the studies by Conley and Glauber (2005) and Black et al. (2005) it remains unclear if unbiased causal estimates of the effect of sibship size on educational attainment are obtained. In the following sections I present the data used in the paper and motivate why I believe that my approach is theoretically and empirically superior to previous cross-sectional and IV studies.

Data and Variables

Data

Data for this study comes from the Wisconsin Longitudinal Study (WLS). The WLS is a longitudinal study of a random sample of 10,317 men and women who graduated from Wisconsin high schools in 1957. 72 percent of the respondents were born in 1939, while the remaining 28 percent were born in adjacent years. Interviews with the respondents or their parents have been carried out in 1957, 1964, 1975, 1992/1993, and 2004. Response rates have remained remarkably high throughout the study period, with around 90 percent of the sample being re-interviewed in the 1964 and 1975 waves (see Sewell and Hauser 1980; Hauser and Sewell 1985; Warren et al. 2002 for more detailed information on the WLS).

The WLS was chosen for this research because of its extremely rich background information on the primary respondents, their parents, as well as one randomly selected child (i.e. three generations of respondents). In addition, the WLS data has background information on the primary respondent's *spouse*, a unique feature which I exploit in this analysis. Notably, the fact that I have background information on the sizes of the sibships from which *both* the primary respondent and his/her spouse originate, as well as their ages at the birth of their first child, is particularly attractive. In this paper I use a sub sample of the WLS respondents and their spouses. In order to be included in the sample the following requirements have to hold: (1) the respondent is continually married throughout the 1964-1975 period (this restriction was necessary because information on spouse's sibship size and age at birth of first child was only available for primary respondents' (reported in 1975) "current spouse"), (2) respondents' children were at least 25 years old by the time of the 1992/1993 interview (this restriction was applied to be reasonably sure that the children had completed their education). These restrictions yield a gross sample of 4,782 respondents.

While the WLS data is well-suited for my research agenda it also has several limitations. First, being comprised of only high-school graduates, the WLS has an under representation of respondents from lower socioeconomic strata who did not attend high school or who dropped out.

This means that the WLS graduates are more homogeneous in terms of socioeconomic characteristics compared to similar cohorts in the total US population. Second, there are only very few African American, Hispanic, or Asian respondents in the WLS. As a consequence, racial differences in educational attainment cannot be analyzed with the WLS.

TABLE 2 ABOUT HERE

Variables

The means and standard deviations of all variables used in the analysis are shown in table 2. The dependent variable in this study is the respondent's child's educational attainment measured by years of completed schooling. As is seen from the table, the mean number of years of schooling completed by the children is 14.11 (SD = 2.32). Information on the child's final educational attainment was provided by the respondents in the 1992/1993 survey. This variable was top-coded at 20 years of schooling (as were all variables measuring educational attainment).

The primary explanatory variable in the analysis is *sibship size*. As is customary in the literature, this variable counts the selected child's total number of brothers and sisters. The highest number of siblings observed in the data is 13, while the mean number of siblings in the WLS sample is 2.30 (SD = 1.43). Other sibship characteristics included in the analysis are birth order, birth spacing, and sibship sex composition. First, *birth order* is measured through a three-category ordinal variable indicating if the selected child is the youngest or oldest sibling, or, alternatively, if the child is in a middle position in the sibship. *Birth spacing* measures the interval in months from the selected child to the next (older or younger) sibling (where twins are coded as 0 months). Finally, *sibship sex* *composition* is an indicator of the relative share of boys (as opposed to girls) in the sibship (with a range from 0 to 1).

Furthermore, in the analysis I include several measures of parents' socioeconomic status. First, *father* and *mother's education* measured in years of schooling is included (again, these variables were top-coded at 20 years). Second, *father's socioeconomic position*, here converted into Duncan's (1961) Occupational Prestige or SEI scores for the longest held job in 1974, was included. It was not possible to include a similar variable for mothers because many mothers in the WLS sample (approximately 40 percent) were not active in the labor market. Third, the natural logarithm of the total earnings of the respondent and his/her spouse in 1974 measured in hundreds of US dollars is included. Fourth, the primary respondent's (either father or mother's) test score on the Henmon-Nelson Test of Mental Ability in 1957 (conducted at approximately age 18) is controlled (see Warren et al. 2002:440-41 for more information on this test). This variable, which is measured prior to respondents acquiring their full education, was included in order to control for intergenerational transfers of mental ability from parents to children (Plug and Vijverberg 2003), but also to control for potential confounding with sibship size arising from low-IQ parents having more children than high-IQ parents. Furthermore, also the sex (coded as a dummy variable with male = 1) and age of the child in years are included as controls. Finally, in addition to sibship, socioeconomic, and demographic variables, four instrumental variables are also included in the analysis. As described above, these variables are the *size of the sibship (i.e. the total number of brothers and sisters) in which the mother and the father grew up*, as well as the *age (in years) of the mother and father at the birth of their first child*. The motivation for using these variables as instruments is described in the next section.

Empirical Strategy

The research questions calls for a statistical model which explains children's educational attainment as a function of three types of variables: sibship size, other sibship characteristics, and socioeconomic and demographic background characteristics. Formulating this model as an Ordinary Least Squares (OLS) regression model, for child *i* (*i = 1,…,n*) I get

$$y_i = \alpha_1 + \beta_1 x_i + \beta_2 ' k_i + \beta_3 ' s_i + \varepsilon_{1i}, \text{ (1)}$$

where *y* is years of completed schooling, α_1 is a constant, *x* is the size of the sibship with regression coefficient β_1, *k* is the vector of additional sibship characteristics (birth order, birth spacing, and sibship sex composition) with coefficient vector $\beta_2 '$, *s* are the socioeconomic background variables and the child's demographic characteristics with regression coefficients $\beta_2 '$, and finally ε_{1i} is a random error term assumed to have a normal distribution and constant variance ($\varepsilon_1 \sim N(0,\sigma^2_{\varepsilon_1})$) (e.g. Greene 2003).

Most empirical studies on sibship size and educational attainment (cf. table 1) interpret the parameter of sibship size, β_1, as the average causal effect of increasing sibship size by one child on the years of schooling completed by child *i*. As shown in table 1, the vast majority of studies obtain a statistically significant negative estimate of β_1. However, in order for the effect of sibship size to be an unbiased causal estimate of the effect of sibship size several assumptions must hold. Most importantly, sibship size must be completely exogenous in equation 1. Exogeneity implies that the variable should exclusively capture the causal effect of sibship size on educational attainment and not the effect of other unobserved or omitted variables summarized in the model error term.

Formally, I write this condition as $E[\varepsilon_1 | x_i] = 0$; i.e. that the expected value of ε_1 given *x* is 0.[5] If indeed sibship size is somehow picking up unobserved socioeconomic or familial characteristics not properly controlled in the model (i.e. $E[\varepsilon_1 | x_i] \neq 0$), its estimated effect on educational attainment β_1 will be biased. The magnitude of this bias depends on the correlation between *x* and ε_1, but generally larger correlations lead to more biased estimates of β_1 and its standard error.

Instrumental Variable Estimation

The method of Instrumental Variables (IV) was designed to deal with the problem of endogenous explanatory variables (e.g. Angrist 1993; Angrist et al. 1996). Suppose that one can find a variable or vector of *instrumental* variables, here denoted *z*, which is correlated with sibship size but which has no direct effect on children's educational attainment (other than the effect acting through sibship size). In this case one may handle the potential endogeneity problem in equation 1 by introducing the additional regression model

$$x_i = \alpha_2 + \lambda' z_i + \beta_4' k_i + \beta_5' s_i + \varepsilon_{2i}, \quad (2)$$

where the dependent variable *x* is sibship size, *z* is a vector comprising the instrumental variables, the *k* and *s* vectors are the same as described above, α_2 is a constant, and ε_{2i} is a normally distributed error term ($\varepsilon_2 \sim N(0, \sigma^2_{\varepsilon_2})$). To clarify the assumptions underlying the IV method observe (1) that in order to identify the model there must be *at least* as many instrumental variables as there are endogenous explanatory variables (i.e. $z \geq x$), (2) that the covariance between *z* and *x* must be different from 0 (i.e. $\text{cov}(z, x) \neq 0$), and (3) that the instrumental variables *z* are assumed to be independent of both error terms (i.e. $E[\varepsilon_1 | z] = 0$ and $E[\varepsilon_2 | z] = 0$). If these assumptions hold, IV estimation may be used, first, to regress sibship size on the vector of instrumental variables, and, second, to plug in the predicted values of sibship size $\tilde{x}$ (instead of *x*) into equation 1, thereby obtaining a consistent causal estimate of sibship size on children's educational attainment (since, by definition, $\tilde{x}$ will not correlated with the error term in equation 1). A range of specification tests devised to test if the assumptions of the IV method are met in empirical applications have been developed. These tests are described below where appropriate.

Selection of Instrumental Variables

Since the IV model is identified using only the portion of variation in sibship size predicted by the instruments (and uncorrelated with the model error term), the interpretation of *how* and for *whom* the estimated causal effect applies is crucial. For example, if in an IV application the endogenous variable is a binary measure of participation in some type of non-experimental treatment to which selection is not random, the estimated effect represents the "local average treatment effect" for those who participated in the treatment (and, hence, the estimate does not pertain to those who were *not* exposed to the treatment) (e.g. Winship and Morgan 1999; DiPrete and Gangl 2004).

In recent years, researchers have increasingly relied on "natural experiments" as a means of mimicking experimental conditions and introducing exogenous variation in IV applications (for reviews see Rosenzweig and Wolpin 2000; Angrist and Krueger 2001). Examples of the use of "natural" random variation in social stratification studies include sibling sex composition (Butcher and Case 1994; Conley and Glauber 2005) and twin births (Rosenzweig and Wolpin 1980; Black et al. 2005). As described earlier, in this analysis I follow the logic of "natural" experiments and use random variation in WLS families' reproductive capabilities as instruments for sibship size. Existing demographic research indicates that a sizeable intergenerational correlation in fertility behavior exits (see Johnson and Stokes 1976; Michael and Tuma 1985; Axinn et al. 1994; Murphy and Wang 2001; Murphy and Knudsen 2002). This correlation is considered to arise from a combination of (1) inherited genetic predispositions and (2) cultural and religious norms regarding the “appropriate” size of the family. The WLS is well-suited for this type of approach because it contains information, first, on the number of siblings in both the mother and father’s family of origin, and, second, the age of both the mother and father at the birth of their first child. As a consequence, in the WLS I have multiple indicators of the reproductive “quality” and normative predispositions of both mothers and fathers. These variables could arguably be used as instruments for sibship size.

In terms of interpretation, my instruments also build on somewhat less restrictive behavioral assumptions than is typically the case in the literature. For example, studies using sibship sex composition as instruments (e.g. Butcher and Case 1994; Conley and Glauber 2005) typically (implicitly or explicitly) assume that parents tend to favor a certain sex compositions (and that sons are more “productive” than daughters), or, at least, that parents somehow adjust their “investments” in children (which in turn affect their intellectual and educational performance) depending on the sex composition of the sibship (see Dahl and Moretti 2004). Studies using twin births as instruments may be less susceptible to such behavioral assumptions, but since twin births are comparatively rare events (less than 3 percent of all births result in twins or more children) this type of instrument usually provides very little exogenous variation in sibship size (which in turn may imply poor identification of the IV model). The instruments used in this study do not require such underlying behavioral assumptions. Rather, they rest on the assertions that the WLS families are comprised from individuals with different reproductive capabilities, and, furthermore, that the WLS parents were unaware of the reproductive capability of their partner when they formed the marriage.

Results

This section presents the results of the empirical analysis of the Resource Dilution Hypothesis claiming a negative causal effect of sibship size on children's educational attainment. In order to illustrate the empirical findings, I estimate four models. These are three OLS models in which, in addition to sibship size, I successively add more explanatory variables to the model. Finally, I estimate the IV model presented above that handles potential endogeneity with respect to sibship size and children's educational attainment.

TABLE 3 ABOUT HERE

The results of the empirical analyses are presented in table 3. OLS I represents the most simple case, as sibship size is the only explanatory variable in the model. Not surprisingly, I find sibship size to have a highly significant (t = -8.18) negative effect on children's educational attainment. The estimate of -.201 is very similar to the effect sizes reported in previous studies (see table 1) and in accordance with the Resource Dilution Hypothesis. In OLS II I include the other sibship variables, i.e. birth order, birth spacing, and sibship sex composition. Interestingly, adding these variables to the model only decreases the negative effect of sibship size marginally to -.199 (with t = -7.03). Among the other sibship variables, only birth order has a significant impact on children's educational attainment. Similar to previous studies (see Cicirelli 1978; Ernst and Angst 1983; Heer 1985), I find that being the oldest sibling (i.e. the first-born child in the family) has a positive effect ($p < .10$) on years of schooling completed, whereas being last-born has a distinct negative effect. This result then supports the notion that first-borns have a comparative advantage over later-born siblings in terms of the amount of parental resources received.

In OLS III I include the socioeconomic and demographic control variables. In this model the effect of sibship size on educational attainment drops considerably to -.103 but remains highly significant (t = -3.96). This finding might indicate that in OLS II the effect of sibship size may be biased due to sibship size picking up some of the effect of the socioeconomic variables present in OLS III. In OLS III birth order is still significant, with first-borns enjoying a significant educational advantage, and last-born experiencing a disadvantage. Finally, the effects of the socioeconomic variables are largely identical to those reported in previous studies. Both father and mother's level of education are positively related to children's educational attainment, as is also father's socioeconomic status. Family income is not significant. Furthermore, the child's sex is not significant, but I find that the older the child the less education they attain. Finally, as expected I find that the higher the IQ score of the primary respondent the more education their children obtain.

By and large, the OLS models I-III reproduce the findings reported in the existing literature. However, the question remains if the negative effect of sibship size on educational attainment reported in OLS III may be interpreted as a true causal effect. In the model labeled *IV* also shown in table 3, I apply the IV estimation method. In the model I use the size of father and mother's sibship, as well as the age of the father and mother at the birth of their first child, as instruments for sibship size. The most important finding arising from the IV model is that the negative effect of sibship size on children's educational attainment is much *stronger* in *IV* compared to in OLS III. In fact, the estimated negative effect of sibship size on years of completed schooling has increased significantly from around one-tenth of a year in OLS III (-.103) to almost one-third of a year (-.305) in *IV*. This finding then supports the Resource Dilution Hypothesis claiming a negative causal effect of sibship size on educational attainment. On the other hand, my results are in contrast to those reported by

Guo and VanWey (1999a), Black et al. (2005) and Conley and Glauber (2005) suggesting that no causal relationship exists between sibship size and educational attainment. Note that the standard error of the estimated effect of sibship size has increased significantly (from .026 to .130). This is because the IV estimator only uses the variation in sibship size which derives from the instrumental variables and which is exogenous to children's educational attainment (the assumption of exogeneity is tested below). However, the negative IV estimate of the effect of sibship size on educational attainment is still significant at the conventional 95 percent level.

But why do the results from OLS III and *IV* differ? Selection into parenthood on unobserved characteristics may play a key role. If, for example, in OLS III sibship size is correlated with unobserved parental characteristics explaining why some parents choose to have many children, and if these parents possess (or develop) certain resources which enable a better or more efficient "management" of large sibships (i.e. the "penalty" for large sibship size implied by the Resource Dilution Hypothesis is less severe in these families compared to in "average" families), then OLS III will underestimate the true *average* negative effect of sibship size on children's educational attainment. By definition, the IV estimate of -.305 represents the (endogeneity-corrected) "pure" *average* causal effect of sibship size for an individual drawn randomly from the sample. This means that for "average" families in the WLS data large sibships do seem to lead to children acquiring less education than families with few children, as claimed by the Resource Dilution Hypothesis.

But is my IV estimate trustworthy? Unlike previous IV studies, the availability of multiple instruments means I am able to examine the efficiency and validity of the IV estimates. As discussed previously, several assumptions have to hold in order for the IV method to provide efficient and unbiased estimates. The first assumption is that the instrumental variables must predict sibship size. In the lower part of table 3 I report some results from the first stage regressions of parents' sibship sizes and ages at birth of first child on sibship size. As may be seen in the table, all four instruments are highly significant predictors of sibship size. Not surprisingly, the larger the size of fathers and mothers' sibships, the more children they are likely to have. Similarly, the older fathers and mothers were at the birth of their first child the fewer children they have. Interestingly, I find that mothers' characteristics are relatively more important than fathers' characteristics with respect to predicting sibship size. Several measures have been developed in the econometrics literature to evaluate the power and efficiency of instrumental variables. It is well-known that "weak" instruments (i.e. instruments which are only weakly correlated with the endogenous variable) pose considerable problems in IV applications due to poor identification of the model (Staiger and Stock 1997). In table 3 I report two conventional diagnostics of weak instruments: The F-test for excluded instruments in the first stage regressions and the partial R^2 for all excluded instruments. With respect to the F-test for excluded instruments, a conventional rule of thumb in the literature on weak instruments (Staiger and Stock 1997) is that the F-statistic for this test should not be below 10. In my case, the F-statistic has a value of 46.08 with an associated *p*-value of less than .0000, meaning that the power of the instruments is sufficiently high to rule out any concern about weak instruments. The partial R^2 of .0423 which describes the relative explanatory power of the instruments in the first stage regression provide a similar picture (Shea 1997).

The second condition for obtaining valid IV estimates is that the instruments must not be correlated with the model error term. If this condition is not satisfied the IV procedure does not solve the endogeneity problem for which it was designed. When single instruments are used (e.g. Black et al. 2005; Conley and Glauber 2005) the IV model is just-identified (i.e. there are exactly as many instruments as there are endogenous variables) and there is insufficient information in the model to test for any correlation between the instruments and the model error term. However, when several instruments are available for the same endogenous variable, the IV model is overidentified and one may examine the validity of the instruments. A conventionally used test for overidentifying restrictions in IV applications is the Sargan test for overidentifying restrictions (Sargan 1958; Baum et al. 2003). In table 3 I report the results of the Sargan test. The Sargan test statistic has a chi-square distributed value of 4.534 with 3 degrees of freedom, yielding a *p*-value of .2093. This result indicates that I accept the null-hypothesis that the instruments are valid.

Conclusion

The ambition of this paper was to implement a formal test of the main proposition in the Resource Dilution Hypothesis: That sibship size has a negative causal effect on educational outcomes. The vast majority of previous studies in sociology using cross-sectional data analysis find a negative effect of sibship size on intellectual and educational outcomes, be those IQ measures, test scores, or educational attainment, and the conventional wisdom in the literature is that the relationship is indeed causal. However, only very few studies have mounted actual tests of the hypothesized causal relationship between sibship size and educational attainment. In addition, using more sophisticated methods recent critics of the Resource Dilution Hypothesis have argued that the negative relationship between sibship size and educational outcomes reported in the literature represents a methodological artifact rather than a true causal effect (Guo and VanWey 1999a; Black et al. 2005; Conley and Glauber 2005).

In order to test if sibship size has a negative causal effect on educational attainment I employ the method of Instrumental Variables (IV). IV methods have been used in two previous studies (Black et al. 2005; Conley and Glauber 2005), but in this paper I improve upon these studies in two key aspects by (1) introducing instrumental variables which capture random variation in parents' reproductive capability to predict fertility rather than depending on behavioral assumptions of how parents "plan" fertility, and (2) exploit the availability of multiple instrumental variables to test the validity of the IV procedure. Previous IV studies do not test the validity of the instruments meaning that it remains unknown if these studies accomplish their aim of providing unbiased estimates of the causal effect of sibship size on educational attainment.

In the empirical analysis, using data from the Wisconsin Longitudinal Study (WLS) and four instruments for sibship size (mother's sibship size, father's sibship size, mother's age at the birth of the first child, and father's age at the birth of the first child), I test if a causal relationship exits between sibship size and children's educational attainment (measured by years of completed schooling). Controlling for other sibship, demographic, and family socioeconomic characteristics, standard OLS estimation indicates that sibship size has a highly significant negative effect on children's educational attainment in the range of one-tenth of a year of schooling per sibling. However, when implementing the IV procedure I find that the negative effect of sibship size on educational attainment in fact appears much stronger and in the range of about one-third of a year of schooling per sibling. Furthermore, the instrumental variables used in the analysis were found to be both relevant and valid. The empirical analysis then provides strong empirical support for the Resource Dilution Hypothesis suggesting that, on average, children's educational attainment decreases as the size of the sibship increases.

Several limitations in the present study and some potential avenues for future research should be mentioned. First, while the WLS data is well-suited for this research because of the availability of information on mothers and fathers' sibship sizes and their ages at the birth of their first child, then several important limitations in the WLS should be considered. Notably, since the WLS comprises a geographically select sample of respondents who have all completed high school, the data is not fully representative of the US population in terms of geographical and socioeconomic diversity. Also, the restriction of the sample to continuously married respondents means that in this study I am unable to analyze the impact of family disruption on children's educational outcomes (e.g. Teachman 1987; Sandefur et al. 1992; Jonsson and Gähler 1997; Biblarz and Gottainer 2000). Finally, the WLS has only very few non-white respondents. This means that this study is silent on potentially important racial differences with respect to the relationship between sibship size and educational outcomes (e.g. Kuo and Hauser 1995; Roscigno and Ainsworth-Darnell 1999; Cameron and Heckman 2001).

As always, the answers provided in this study prompts new questions. For example, the approach taken here and in most of the literature only captures the effect of sibship size on educational attainment by proxy. How *precisely* does having many brothers and sisters affect educational attainment? The Resource Dilution Hypothesis states that a range of parental resources play a role (money, time, care), but in this study I (and most existing studies) only estimate the somewhat "fuzzy" marginal effect of *total* sibship size on educational attainment. Exactly *how* this effect works remains unresolved. Indeed, as the IV estimates presented here indicate that the negative effect of sibship size on educational attainment is somewhat stronger than previously reported, it becomes highly relevant to disentangle how this effect works. A few previous studies have included explanatory variables pertaining to "softer" interpersonal and social parental resources, thereby attempting to account for more qualitative dimensions of parental resources than those typically studied in the literature (e.g. Downey 1995). However, while certainly significant contributions, this approach – within the framework of IV estimation – risks introducing more potentially endogenous explanatory variables into the analysis.

Finally, marginal effects of sibship size on educational attainment tell little of how parents “invest” resources in children *over time*. A few studies have exploited longitudinal data to investigate the role of longitudinal changes in parental resources and investments in children (e.g. Baydar et al. 1997; Guo and VanWey 1999a). This approach enables a more precise and detailed analysis of the short- and long-term impacts of parental resources on children’s outcomes. Unfortunately, lack of appropriate data has so far prevented further developments of this type of analysis.

Notes

[1] A theoretical contender to the Resource Dilution Hypothesis is the *Confluence Model*. The Confluence Model attributes the negative relationship between sibship size and children's intellectual outcomes to a poor intellectual climates and stimuli in large families (Zajonc and Markus 1975). That is, more siblings "drain" the family's intellectual milieu, which in turn implies that children tend to fare poor intellectually. The Confluence Model is not considered in detail here, first, because it pertains mostly to children's intellectual (and not educational) outcomes, and, second, because most empirical evidence does not support the Confluence Model (for reviews see Heer 1985; Steelman 1985; Steelman et al. 2002).

[2] This review focuses solely on the effect of sibship size on *educational attainment*. For reviews of other outcome variables such as IQ or tests score results see Cicirelli (1978), Ernst and Angst (1983), Heer (1985), Steelman (1985), and Steelman et al. (2002).

[3] Downey et al. (1999; see also Steelman et al. 2002:254-56) suggest that the fixed effect model is *fundamentally* too conservative in its evaluation of the relationship between sibship size and intellectual outcomes, and they use the study by Baydar et al. (1997) which also uses change scores and the same data (the NLSY) (and which finds that the addition of an additional sibling to the family has a deleterious effect on several social-psychological outcome measures) as a "counter-case" to the analysis of Guo and VanWey. However, these objections are not correct for two reasons. First, the fixed effect model adjusts for *all types* of unobserved, time-invariant variables not included in the model irrespective of whether or not these unobserved variables are correlated with other observed variables. These properties arguably make the fixed effect model the most robust of all panel data estimators since it (1) eliminates all time-invariant unobserved heterogeneity, and (2) carries the least number of assumptions (concerning the distribution of the unobservables, correlations between observables and unobservables etc.). This means that Guo and VanWey in fact put the Resource Dilution Hypothesis to the test using an appropriate estimation technique. Obviously, the drawback of the fixed-effect estimator is that it only removes *time-invariant* (and thus not time-varying) heterogeneity, and, furthermore, that it is highly inefficient (since only the variation in *changes* in the dependent and explanatory variables rather than cross-sectional variation in the total sample is used). Both of these drawbacks may have important implications in the *actual* study carried out by Guo and VanWey, but this does not relate to the theoretical properties of the fixed effect estimator per se. Second, the study by Baydar et al. (1997) does not implement a fixed effect model, but rather a standard cross-sectional regression model including change scores in some of the explanatory variables. However, this "hybrid" model

does not possess the same properties as the fixed effect model (notably, it does not correct for unobserved heterogeneity) and should not be compared to the model used by Guo and VanWey.

[4] Strangely, in a similar ordinary logit model Conley and Glauber (2005) find that going from two to three children actually *increases* the likelihood of children attending private school. This result is somewhat anomalous from the perspective of the Resource Dilution Hypothesis.

[5] Obviously, the same assumption applies to all explanatory variables in equation 1 so that the full expectation is in fact $E[\varepsilon_1 \mid x_i, k_i, s_i] = 0$.

References

Alwin, Duane F., and Arland Thornton. 1984. "Family Origins and the Schooling Process: Early Versus Late Influences of Parental Characteristics." *American Sociological Review* 49:784-802.

Angrist, Joshua D. 1993. "The Effect of Veteran Benefits on Education and Earnings." *Industrial and Labor Relations Review* 46(4):637-52.

Angrist, Joshua D., Guido W. Imbens, and Donald B. Rubin. 1996. "Identification of Causal Effects Using Instrumental Variables." *Journal of the American Statistical Association* 91:444-55.

Angrist, Joshua D., and Alan B. Krueger. 2001. "Instrumental Variables and the Search for Identification: From Supply and Demand to Natural Experiments." *Journal of Economic Perspectives* 15(4):69-85.

Axinn, William G., Marin E. Clarkberg, and Arland Thornton. 1994. "Family Influences on Family Size Preferences." *Demography* 31(1):65-79.

Baum, Christopher F., Mark E. Schaffer, and Steven Stillman. 2003. "Instrumental Variables and GMM: Estimation and Testing." Boston College, Department of Economics. Working Paper No. 545.

Baydar, Nazli, Patricia Hyle, and Jeanne Brooks-Gunn. 1997. "A Longitudinal Study of the Effects of the Birth of a Sibling During Preschool and Early Grade School Years." *Journal of Marriage and the Family* 59:957-65.

Becker, Gary S., and Nigel Tomes. 1986. "Human Capital and the Rise and Fall of Families." *Journal of Labor Economics* 4:S1-S39.

Behrman, Jere R., Mark R. Rosenzweig, and Paul Taubman. 1994. “Endowment and the Allocation of Schooling in the Family and in the Marriage Market: The Twins Experiment.” *Journal of Political Economy* 102(6):1131-74.

Biblarz, Timothy J., and Greg Gottainer. 2000. “Family Structure and Children’s Success: A Comparison of Widowed and Divorced Single-Mother Families.” *Journal of Marriage and the Family* 62:533-48.

Black, Sandra E., Paul G. Devereaux, and Kjell G. Salvanes. 2005. “The More the Merrier? The Effect of Family Composition on Children’s Education.” *Quarterly Journal of Economics* 120(2):669-700.

Blake, Judith. 1981. “Family Size and the Quality of Children.” *Demography* 18(4):421-42.

-----. 1985. “Number of Siblings and Educational Mobility.” *American Sociological Review* 50:84-94.

-----. 1989. *Family Size and Achievement.* University of California Press.

Butcher, Kristin F., and Anne Case. 1994. “The Effects of Sibling Sex Composition on Women’s Education and Earnings.” *The Quarterly Journal of Economics* 109(3):531-63.

Cameron, Stephen V., and James J. Heckman. 2001. “The Dynamics of Education Attainment for Black, Hispanic, and White Males.” *Journal of Political Economy* 109:455-99.

Case, Anne, I-Fen Lin, and Sara McLanahan. 2001. “Educational attainment of siblings in stepfamilies.” *Evolution and Human Behavior* 22:269-89.

Cicirelli, V. G. 1978. “The relationship of family structure to intellectual abilities and achievement.” *Review of Educational Research* 48(3):365-79.

Conley, Dalton. 2000. “Sibship Sex Composition: Effects on Educational Attainment.” *Social Science Research* 29:441-57.

-----. 2001. “Capital for College: Parental Assets and Postsecondary Schooling.” *Sociology of Education* 74:59-72.

Conley, Dalton, and Rebecca Glauber. 2005. “Sibship Size, Birth Order and Educational Stratification: IV estimates.” Unpublished paper.

Dahl, Gordon B., and Enrico Moretti. 2004. “The Demand for Sons: Evidence from Divorce, Fertility, and Shotgun Marriage.” *National Bureau of Economic Research Working Paper 10281*.

Datcher, Linda. 1982. “Effects of Community and Family Background on Achievement.” *Review of Economics and Statistics* 64(1):32-41.

Davies, Scott. 1995. “Reproduction and resistance in Canadian high schools: an empirical examination of the Willis Thesis.” *British Journal of Sociology* 46(4):662-87.

de Graaf, Paul M. 1986. “The impact of financial and cultural resources on educational attainment in the Netherlands.” *Sociology of Education* 59:237-46.

-----. 1988. “Parents’ financial and cultural resources, grades, and transitions to secondary school in the Federal Republic of Germany.” *European Sociological Review* 4(3):209-21.

de Graaf, Paul M., and Johannes J. Huinink. 1992. “Trends in Measured and Unmeasured Effects of Family Background on Educational Attainment and Occupational Status in the Federal Republic of Germany.” *Social Science Research* 21:84-112.

DiPrete, Thomas, and Markus Gangl. 2004. “Assessing Bias in the Estimation of Causal Effects: Rosenbaum Bounds on Matching Estimators and Instrumental Variables Estimation with Imperfect Instruments.” *Sociological Methodology* 34:271-310.

Downey, Douglas B. 1995. “When bigger is not better: Family size, parental resources, and children’s educational performance.” *American Sociological Review* 60:746-61.

-----. 2001. “Number of Siblings and Intellectual Development. The Resource Dilution Explanation.” *American Psychologist* 56(6/7):497-504.

Downey, Douglas B., Brian Powell, Lala C. Steelman, and Shana Pribesh. 1999. “Much ado about siblings: Change models, sibship size, and intellectual development.” *American Sociological Review* 64:193-98.

Duncan, Beverly. 1967. “Education and Social Background.” *American Journal of Sociology* 72:363-72.

Duncan, Otis D. 1961. “A Socioeconomic Index for All Occupations.” Pp. 109-138 in *Occupations and Social Status,* edited by A. J. Reiss Jr. Free Press of Glencoe.

Duncan, Otis D., David L. Featherman, and Beverly Duncan. 1972. *Socioeconomic Background and Achievement.* Seminar Press.

Ermisch, John, and Marco Francesconi. 2001. “Family Matters: Impact of Family Background on Educational Attainments.” *Economica* 68:137-56.

Ernst, Cécile, and Jules Angst. 1983. *Birth Order. Its Influence on Personality.* Springer-Verlag.

Evans, M. D. R., Jonathan Kelley, and Richard A. Wanner. 2001. “Educational attainment of the children of divorce: Australia, 1940-90.” *Journal of Sociology* 37(3):275-97.

Featherman, David L., and Robert M. Hauser. 1976. “Sexual Inequalities and Socioeconomic Achievement in the U.S., 1962-1973.” *American Sociological Review* 41:462-83.

-----. 1978. *Opportunity and Change.* Academic Press.

Galbraith, Richard D. 1982. “Sibling spacing and intellectual development: A closer look at the confluence model.” *Development Psychology* 18:151-75.

Greene, William H. 2003. *Econometric Analysis.* Prentice Hall.

Guo, Guang, and Leah K. VanWey. 1999a. “Sibship Size and Intellectual Development: Is the Relationship Causal?” *American Sociological Review* 64:169-87.

-----. 1999b. “The effects of closely spaced and widely spaced sibship size on intellectual development.” *American Sociological Review* 64:199-206.

Hauser, Robert M., and David L. Featherman. 1976. “Equality of schooling: trends and prospects.” *Sociology of Education* 49:99-120.

Hauser, Robert M., and Hsiang-Hui D. Kuo. 1998. “Does the Gender Composition of Sibships Affect Women’s Educational Attainment?” *Journal of Human Resources* 33(3):644-57.

Hauser, Robert M., and William H. Sewell. 1985. “Birth Order and Educational Attainment in Full Sibships.” *American Educational Research Journal* 22(1):1-23.

Haveman, Robert, and Barbara Wolfe. 1995. “The determinants of children’s attainments: a review of methods and findings.” *Journal of Economic Literature* 33:829-78.

Heer, David M. 1985. “Effects of Sibling Number on Child Outcome.” *Annual Review of Sociology* 11:27-47.

Hill, Martha S., and Greg J. Duncan. 1987. “Parental Income and the Socioeconomic Attainment of Children.” *Social Science Research* 16:39-73.

Johnson, Nan E., and Shannon Stokes. 1976. “Family Size in Successive Generations: The Effects of Birth Order, Intergenerational Change in Lifestyle, and Familial Satisfaction.” *Demography* 13(2):175-87.

Jonsson, Jan O., and Michael Gähler. 1997. “Family Dissolution, Family Reconstitution, and Children’s Educational Careers: Recent Evidence for Sweden.” *Demography* 34(2):277-93.

Kaestner, Robert. 1997. “Are Brothers really Better? Sibling Sex Composition and Educational Achievement Revisited.” *Journal of Human Resources* 32(2):250-275.

Krein, Sheila F., and Andrea H. Beller. 1988. “Educational Attainment of Children From Single-Parent Families: Differences by Exposure, Gender and Race.” *Demography* 25(2):221-34.

Kuo, Hsiang-Hui D., and Robert M. Hauser. 1995. “Trends in Family Effects on the Education of Black and White Brothers.” *Sociology of Education* 68:136-60.

-----. 1997. “How Does Size of Sibship Matter? Family Configuration and Family Effects on Educational Attainment.” *Social Science Research* 26:69-94.

Lindert, Peter H. 1977. “Sibling Position and Achievement.” *Journal of Human Resources* 12:198-219.

Lucas, Samuel R. 2001. “Effectively Maintained Inequality: Education Transitions, Track Mobility, and Social Background Effects.” *American Journal of Sociology* 106:1642-90.

Mare, Robert D. 1980. “Social Background and School Continuation Decisions.” *Journal of the American Statistical Association* 75:295-305.

Mare, Robert D., and Meei-Shenn Tzeng. 1989. “Fathers’ Ages and the Social Stratification of Sons.” *American Journal of Sociology* 95:108-31.

Marjoribanks, Kevin, and Herbert J. Walberg. 1975. “Birth Order, Family Size, Social Class and Intelligence.” *Social Biology* 22:261-68.

Michael, Robert T., and Nancy B. Tuma. 1985. “Entry into Marriage and Parenthood by Young Men and Women: The Influence of Family Background.” *Demography* 22(4):515-44.

Micklewright, John. 1989. “Choice at Sixteen.” *Economica* 56:25-39.

Murphy, Mike, and Lisbeth B. Knudsen. 2002. “The intergenerational transmission of fertility in contemporary Denmark: The effects of number of siblings (full and half), birth order, and whether male or female.” *Population Studies* 56:235-48.

Murphy, Mike, and D. Wang. 2001. “Family-level Continuities in Childrearing in Low-Fertility Societies.” *European Journal of Population* 17:75-96.

Olneck, Michael R., and David B. Bills. 1979. “Family Configuration and Achievement: Effects of Birth Order and Family Size in a Sample of Brothers.” *Social Psychology Quarterly* 42:135-48.

Phillips, Meredith. 1999. “Sibship size and academic achievement: What we know and what we still need to know.” *American Sociological Review* 64:188-92.

Plug, Erik, and Wim Vijverberg. 2003. “Schooling, Family Background, and Adoption: Is it Nature or Nurture?” *Journal of Political Economy* 111(3):611-41.

Pong, Suet-Ling. 1997. “Sibship Size and Educational Attainment in Peninsular Malaysia: Do Policies Matter?” *Sociological Perspectives* 40(2):227-42.

Powell, Brian, and Lala C. Steelman. 1990. “Beyond Sibship Size: Sibling Density, Sex Composition, and Educational Outcomes.” *Social Forces* 69(1):181-206.

-----. 1993. “The Educational Benefits of Being Spaced Out: Sibship Density and Educational Progress.” *American Sociological Review* 58:367-81.

Ribar, David C. 1991. “A Multinomial Logit Analysis of Teenage Fertility and High School Completion.” *Economics of Education Review* 12(2):153-64.

Rodgers, Joseph L. 2001. “What causes birth order-intelligence patterns.” *American Psychologist* 56:505-10.

Roscigno, Vincent J., and James W. Ainsworth-Darnell. 1999. “Race, Cultural Capital, and Educational Resources: Persistent Inequalities and Achievement Returns.” *Sociology of Education* 72:158-78.

Rosenzweig, Mark R. 1986. “Birth Spacing and Sibling Inequality: Asymmetric Information Within the Family.” *International Economic Review* 27(1):55-76.

Rosenzweig, Mark R., and Kenneth Wolpin. 1980. “Testing the Quality-Quantity Fertility Model: The Use of Twins as a Natural Experiment.” *Econometrica* 48(1):227-40.

-----. 2000. “Natural ‘Natural’ Experiments in Economics.” *Journal of Economic Literature* 38(4):827-74.

Sandefur, Gary, Sara McLanahan, and Roger A. Wojtkiewicz. 1992. “The Effects of Parental Marital Status during Adolescence on High School Graduation.” *Social Forces* 71:103-21.

Sandefur, Gary, Ann Meier, and Mary E. Campbell. In press. “Family resources, social capital, and college attendance.” *Social Science Research*.

Sandefur, Gary, and Thomas Wells. 1999. “Does Family Structure Really Influence Educational Attainment?” *Social Science Research* 28:331-57.

Sargan, John D. 1958. “The Estimation of Economic Relationships Using Instrumental Variables.” *Econometrica* 26:393-415.

Sewell, William H., and Robert M. Hauser. 1980. “The Wisconsin Longitudinal Study of Social and Psychological Factors in Aspirations and Achievements.” *Research in Sociology of Education and Socialization* 1:59-99.

Shavit, Yossi, and Jennifer L. Pierce. 1991. “Sibship Size and Educational Attainment in Nuclear and Extended Families: Arabs and Jews in Israel.” *American Sociological Review* 56:321-30.

Shea, John. 1997. “Instrument Relevance in Multivariate Linear Models: A Simple Measure.” *Review of Economics and Statistics* 79(2):348-52.

Sieben, Inge, Johannes Huinink, and Paul M. de Graaf. 2001. “Family Background and Sibling Resemblance in Educational Attainment. Trends in the Former FRG, the Former GDR, and the Netherlands.” *European Sociological Review* 17(4):401-30.

Staiger, Douglas, and James H. Stock. 1997. “Instrumental Variables Regression with Weak Instruments.” *Econometrica* 65(3):557-86.

Steelman, Lala C. 1985. “A tale of two variables: a review of the intellectual consequences of sibship size and birth order.” *Review of Educational Research* 55:353-86.

Steelman, Lala C., Brian Powell, Regina Werum, and Scott Carter. 2002. “Reconsidering the Effects of Sibling Configuration: Recent Advances and Challenges.” *Annual Review of Sociology* 28:243-69.

Sun, Yongmin. 1999. “The Contextual Effects of Community Social Capital on Academic Performance.” *Social Science Research* 28:403-26.

Teachman, Jay D. 1987. “Family Background, Educational Resources, and Educational Attainment.” *American Sociological Review* 52:548-57.

Warren, John R., Robert M. Hauser, and Jennifer T. Sheridan. 2002. “Occupational Stratification Across the Life Course: Evidence from the Wisconsin Longitudinal Study.” *American Sociological Review* 67:432-55.

Winship, Christopher, and Stephen L. Morgan. 1999. “The Estimation of Causal Effects from Observational Data.” *Annual Review of Sociology* 25:659-707.

Zajonc, Robert B., and Gregory B. Markus. 1975. “Birth Order and Intellectual Development.” *Psychological Review* 82(1):74-88.

Table 1. Summary of the Effect of Sibship Size on Years of Completed Schooling in Previous Studies

Study	Estimate of Sibship Size	Data	Nationality of Sample
Duncan 1967	-.14 to -.22[a,d,m]; -.04 to -.13[c,d,m]	OCG, 1962	US
Duncan et al. 1972	-.21[a,m]	OCG, 1962	US
Featherman and Hauser 1976	-.287[m(OCG62)]; -.202[w(OCG62)]; -.291[m(OCG73)]; -.214[w(OCG73)]	OCG, 1962 and 1973, married respondents	US
Hauser and Featherman 1976	-.19 to -.26[m,d]	OCG, 1962 and 1973	US
Featherman and Hauser 1978	-.184 to -.227[m,d]	OCG, 1973	US
Olneck and Bills 1979	-.153[m,s,e]	OCG, 1962 and 1973; Kalamazoo Brothers	US
Blake 1981	Negative[f]	Various	US
Datcher 1982	-.255[a]; NS[b]	PSID	US
Alwin and Thornton 1984	-.16[a]	Sample of white families in Detroit, Michigan, 1961	US
Blake 1985	Negative[f,m]	OCG, 1962 and 1973, GSS, 1972-1983	US
Hauser and Sewell 1985	Negative[f]	WLS	US
de Graaf 1986	-.112[s]	1977 Quality of Life Survey	Netherlands
Hill and Duncan 1987	-.19[m]; -.14[w]	PSID	US
Teachman 1987	-.112[a,m]; -.119[a,w]	NLS	US
Krein and Beller 1988	-.19[a,m]; -.13[b,m]; -.14[b,w]	NLS	US
Blake 1989	-.199 to -.240[a,m]; -.171 to -.193[a,w]	OCG, 1962 and 1973, GSS, 1972-1986	US
Mare and Tzeng 1989	-.168[m]	OCG, 1973	US
Shavit and Pierce 1991	Negative (Jews)[f,m]; NS (Arab men)	Representative Israeli Jewish/Arab sample	Israel
de Graaf and Huinink 1992	-.066 to -.153[d,s]	German Life History Study	Germany
Butcher and Case 1994	-.507[a,m,e]; -.186[a,w,e]	PSID	US
Hauser and Kuo 1998	Negative [f,w]	OCG, 1973; SFH, 1986/1988; NSFH, 1989	US
Sandefur and Wells 1999	-.10[s]	NLSY	US
Conley 2000	-.10	PSID	US
Case et al. 2001	-.057	PSID	US
Conley 2001	-.124	PSID	US
Evans et al. 2001	-.12	International Social Science Surveys	Australia
Sieben et al. 2001	-.081 to -.236 (FRG)[s];	Various data sets	(Former)

	-.089 to -.178 (GDR)[s]; -.029 to -.03 (Netherlands)[s]		Federal Republic of Germany; (Former) German Democratic Republic; Netherlands
Plug and Vijverberg 2003	-.152	WLS	US

Note. [a] Whites, [b] Blacks, [c] Non-whites, [m] Men, [w] Women, [s] Sibling model, [d] Estimate varies by cohort, [e] Includes non-linear effect of sibship size, [f] Estimate not presented in metric scale. NS = No significant effect. Estimates are shown for models controlling the maximum number of other socioeconomic, familial, and demographic variables available. Results apply to both men and women unless stated otherwise. Data sets: OCG = Occupational Change in a Generation, PSID = Panel Study of Income Dynamics, WLS = Wisconsin Longitudinal Study, GSS = General Social Survey, NLS = National Longitudinal Study, NLSY = National Longitudinal Survey of Youth, GSOEP = German Socioeconomic Panel.

Table 2. Descriptive Statistics. Means and Standard Deviations

Variable	Mean	SD
Child's years of education	14.11	2.32
Child's sex (= male)	.43	.49
Child's age	28.77	2.44
Number of siblings	2.30	1.43
Birth order:		
Oldest sibling	.42	.49
Youngest sibling	.17	.37
Middle sibling	.41	.50
Birth spacing[a]	26.08	26.58
Sex composition[b]	.48	.33
Father's years of education	13.35	2.49
Mother's years of education	12.80	1.73
Father's socioeconomic status	47.82	24.57
Log family income	4.70	1.44
Parent's Henmon-Nelson IQ score	99.89	14.50
Father's sibship size	3.35	2.72
Mother's sibship size	3.35	2.65
Father's age at birth of first child	23.91	3.19
Mother's age at birth of first child	21.79	2.54

Note. N = 4,782, [a] Number of months between birth of selected child and birth of next youngest/next oldest sibling, [b] Number of boys in sibship as proportion of total siblings.

Table 3. Results of OLS and Instrumental Variable Regressions of Children's Educational Attainment. Standard Errors in Parenthesis and *T*-values in Brackets

Model	*OLS I*	*OLS II*	*OLS III*	*IV*
Number of siblings	-.201***	-.199***	-.103***	-.305*
	(.025)	(.028)	(.026)	(.130)
	[-8.18]	[-7.03]	[-3.96]	[-2.40]
Birth order:				
Oldest sibling		.162†	.153†	-.133
		(.088)	(.083)	(.194)
Youngest sibling		-.216*	-.207*	-.514*
		(.112)	(.102)	(.216)
Birth spacing		.001	-.001	-.001
		(.002)	(.002)	(.002)
Sex composition		.079	.014	.104
		(.117)	(.133)	(.145)
Child's sex (= male)			.088	.122
			(.080)	(.083)
Child's age			-.061***	-.047**
			(.015)	(.017)
Father's years of education			.200***	.201***
			(.017)	(.017)
Mother's years of education			.208***	.202***
			(.022)	(.023)
Father's socioeconomic status			.009***	.008***
			(.002)	(.001)
Log family income			.020	.014
			(.023)	(.024)
Parent's Henmon-Nelson IQ score			.012***	.013***
			(.002)	(.002)
Constant	14.573***	14.455***	8.901***	9.213***
	(.067)	(.135)	(.581)	(.613)
Adjusted R^2	.015	.018	.207	-
N	4,253	4,252	4,187	4,186
Selected Results from First Stage IV Regressions of Sibship Size				
Father's sibship size				.019**
				(.007)
Mother's sibship size				.048***
				(.007)
Father's age at birth of first child				-.031***
				(.008)
Mother's age at birth of first child				-.090***
				(.010)
IV Diagnostics				
P-value for F-test for excluded instruments in first-stage regression (F = 46.08)				.0000
Partial R^2 for all excluded instruments				.0423
P-value for Sargan test ($\chi^2_{(3)} = 4.534$)				.2093

Note. *** $p < .001$, ** $p < .01$, * $p < .05$, † $p < .10$ (two-tailed tests)

Appendix Table 1. Summary of the Effect of Sibship Size on Educational Attainment in Studies using Probability Models or with Test Scores/GPA as Outcome Variables

Study	Estimate of Sibship Size	Data	Nationality of Sample	Outcome Variable
		Probability models		
Lindert 1977	Negative[s]	The New Jersey Sibling Sample	US	Ordered indicator of educational attainment
Mare 1980	Negative[m]	OCG, 1973; CPS, 1964	US	Series of 6 educational transitions
de Graaf 1988	Negative	German data set	Germany	Educational choice at age 10: (1) extended elementary school, (2) middle school, (3) gymnasium
Micklewright 1989	Positive (i.e. higher probability of leaving school)	NCDS	Great Britain	Probability of leaving school at minimum permissible age
Ribar 1991	Negative[w]	NLSY	US	Probability of having graduated from high school at age 20
Powell and Steelman 1993	(1) Positive (2) Negative	HSB	US	Probability of (1) High-school drop-out; (2) Attended post-secondary schooling
Behrman et al. 1994	Negative[a]	NLS	US	Postsecondary attainment: (1) None, (2) Two-year degree, (3) Four-year degree
Davies 1995	Negative	Adolescent Experience Survey	Canada	Dummy for planning to go to university (= 1) after secondary school
Jonsson and Gähler 1997	(1) Negative (2) Negative/almost NS	Swedish register data	Sweden	Educational transitions: probability of (1) continuing to Gymnasium; (2) reaching upper secondary school
Pong 1997	Negative	Second Malaysian Life Survey	Malaysia	Proportion having completed secondary education
Ermisch and Francesconi 2001	Negative	BHPS	Great Britain	Educational attainment, 7 ordered categories
Lucas 2001	Negative	HSB	US	Educational transitions: probability of continuing to (1) grade

				11, (2) grade 12, (3) college
Sandefur et al. in press	Negative	NELS	US	Post secondary attainment: (1) None, (2) Certificate, (3) Two-year degree, (4) Four-year degree
		Test Score Results/GPA		
Powell and Steelman 1990	Negative	(1) HSB, (2) NLS	US	Self-reported grades, verbal and math test scores
Downey 1995	Negative[e]	NELS	US	(1) Self-reported grades, (2) standardized math test score, (3) standardized reading tests score
Baydar et al. 1997	(1) Negative effect for children from economically disadvantaged groups (interaction); (2) NS	NLSY	US	Cognitive test score: (PIAT) (1) Reading: recognition, (2) Reading comprehension
Guo and VanWey 1999a	NS	NLSY	US	Cognitive test scores: (1) PPVT, (2) PIAT-R, (3) PIAT-M.
Roscigno and Ainsworth-Darnell 1999	Negative	NELS	US	GPA, mathematics-reading composite test score
Sun 1999	Negative	NELS	US	Test scores in (1) Science, (2) math, (3) reading, (4) social studies

Note: [a] Whites, [b] Blacks, [c] Non-whites, [m] Men, [w] Women, [s] Sibling model, [d] Estimate varies by cohort, [e] Includes non-linear effect of sibship size, [f] Estimate not calculated/presented. NS = No significant effect. Results are shown for models controlling the maximum number of other socioeconomic, familial, and demographic variables available. Results apply to both men and women unless stated otherwise. Data sets: OCG = Occupational Change in a Generation, PSID = Panel Study of Income Dynamics, WLS = Wisconsin Longitudinal Study, GSS = General Social Survey, NLS = National Longitudinal Study, NLSY = National Longitudinal Survey of Youth, NELS = National Education Longitudinal Study, SFH = Survey of Families and Households, NSFH = National Survey of Families and Households, CPS = Current Population Survey, HSB = High School and Beyond, GSOEP = German Socioeconomic Panel, NCDS = National Child Development Study, BHPS = British Panel Household Study.