THE FAMILY DEMOGRAPHY OF HIGHER EDUCATION

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Michael W. Spiller

January 2014

THE FAMILY DEMOGRAPHY OF HIGHER EDUCATION

Michael W. Spiller, Ph.D.
Cornell University 2014

Patterns of educational attainment in the United States have changed over the 20th century, with a significant increase in the value of and demand for college education since the 1980s. Simultaneously, the size of families shrank and the proportion of youth living in two-parent "traditional" households decreased, leading to a proliferation of new family forms. Social scientists have long investigated the relationship between family structure and educational attainment. This dissertation contributes to prior research on families and education by examining the relationship between family structure and enrollment in and completion of 4-year college. The first chapter of the dissertation analyzes two panels of the National Longitudinal Survey of Youth (NLSY) to determine whether the relationship between family size and higher educational attainment changed between the birth cohort completing high school in the early 1980s and the one completing high school in the late 1990s. It also examines whether family income plays a role in determining whether family size impacts higher educational attainment. The second chapter analyzes the later panel of the NLSY to evaluate competing explanations for the negative relationship between family size and educational attainment. Additionally, it examines whether the relationship varies by youths' race/ethnicity. The final chapter presents a measure of family structure that combines the number of family transitions a youth has experienced and a qualitative measure of family type. It then uses propensity score models to examine whether the

negative relationship between non-traditional family structures and higher educational attainment is causal in the later panel of the NLSY. The first chapter finds that there is a negative relationship between family size and higher educational attainment among both birth cohorts. However, it finds that the relationship is concentrated among higher income families in the early panel and lower income families in the later panel. This shift over time is likely due to large changes in higher education aid policies such as the introduction of unsubsidized Stafford loans in 1993. The second chapter finds little support for three explanations claiming that the relationship between family size and higher education is not causal or for the claim that the relationship operates via decreased intellectual ability. It also finds that there is variation in the relationship between family size and higher education by race/ethnicity, with no detectable relationship for Hispanic youth. The final chapter finds that there is a significant causal relationship between being raised in a non-traditional family structure and higher education. Additionally, it finds that the strength of the relationship varies by the likelihood of having a non-traditional family, with the effects concentrated among those who are least likely to have one. This may indicate that communities in which non-traditional families are common provide resources that moderate the impact of non-traditional family structures on educational attainment.

BIOGRAPHICAL SKETCH

Michael Spiller was born and raised in Wichita Falls, Texas. He completed his B.S. in sociology at Midwestern State University in Wichita Falls, for which he was awarded the President's Medal of Honor. He then attended Cornell University where he researched respondent-driven sampling for his M.A. in sociology. After working on a multi-year research project surveying low-wage workers to identify violations of labor and employment laws, he turned to research in family demography for his dissertation. He is currently employed as an epidemiologist by the U.S. Centers for Disease Control and Prevention.

Michael's academic interests include demography, causal inference, sampling methodology, statistics, social network analysis, and epidemiology. He also enjoys programming in R and using data to improve the way things work. In his leisure time, he likes playing the drums and guitar, cooking, and reading the New Yorker from cover to cover. He has spent his life looking for answers; if you have any, please let him know.

For Sally

ACKNOWLEDGMENTS

I am extremely grateful for the guidance of my dissertation advisor, Steve Morgan, and the support of my committee members Kim Weeden, Dan Lichter, and Doug Heckathorn. My deepest thanks go out to them for their advice, encouragement, and service as models for aspiring social scientists.

I have immense appreciation for Sue Meyer, Betsy Collins, Sharon Sandlan, and Alice Murdock, who provided amazing assistance with navigating the red tape of university bureaucracy and the complexities of grant application and management. I would also like to thank Marty White, who was always willing to help punish misbehaving technology and to share his hard-earned wisdom about the ways of the computer.

My years at Cornell were deeply enriched by my friends and peers, with whom I was glad to experience the joys and sorrows of academic life. I will never forget daily lunches with the sociology department staff and students, which provided laughter that fueled my days like coffee never could.

I am deeply appreciative of my family's support and willingness to tolerate me in my youth. Special thanks go out to my grandparents Dorothy and Jim for supporting me in all my endeavors and for instilling in me a lifelong thirst for knowledge.

This dissertation would not have been possible without the incredible, unwavering support of my parents Diane and Mike. I am in awe of their generosity and can only hope to be for my children what they have been for me.

Finally, I am eternally grateful to my wife Sally for taking this journey with me. Her fortitude was immense, as was her patience. Her love carried me through.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

Censuses were the first large scale collection of data about human populations. They allowed rulers to determine approximately how many subjects they had, who needed to be taxed, and how much revenue would be generated by a given tax rate. Censuses also informed forecasts of tax revenue based on the projected size of the population in future years. Projections of population size hinge on rates of fertility and mortality. One measure of fertility is the general fertility rate, the average number of live births per woman of childbearing age. Because most births historically occurred in the context of two-parent families, this number in combination with rates of child mortality provided perhaps the first measure of what is now called family structure: how many people were in the average family.

Prior to industrialization, the primary impact of family size was on a family's agricultural output. Each child added to the family's labor pool and increased its total output as long as he produced more than he consumed. However, this dynamic changed in the United States over the course of the 19th century as the average number of births per woman plunged from about 7 to about 3.5. A number of factors, such as falling infant mortality rates, contributed to this decline, but clearly something was changing in the calculus of whether or not to have additional children. This calculus continued to shift with industrialization and as increasing proportions of the population lived in urban areas. The rise of compulsory schooling for children in the United States in the late 19th century caused another shift in the marginal value of additional children; unless one lived on a family farm, children imposed economic costs with few corresponding economic benefits. (Of course, children have non-economic value, which is presumably why people continue to produce them.)

Along with the size of families, the United States' economy changed significantly over the course of the 19th and 20th centuries. The proportion of the

workforce employed in agriculture declined precipitously from 1850 to 1950, reflecting advances in the mechanization of farming and the rise of industrialization. Service and industrial employment increased over the early 20th century, with a decline in industrial employment beginning around 1960. In the late 20th century, the service sector of the economy grew significantly and, with it, the demand for skilled workers. This increase in demand fueled an increase in the economic returns to education, which began to rise in the 1970s. From the 1980s onward the returns to education, and to higher education in particular, rapidly increased. This led to growing popular demand for and political interest in access to college.

In the mid-20th century, sociologists became increasingly interested in the degree to which children attained higher occupational or socio-economic status than did their parents, or "social mobility." The status attainment literature, in particular, focused on the relationships between fathers' and sons' occupations and the causal processes that produced these relationships. Status attainment scholars recognized that education played a significant role in the production of social mobility. Other sociologists studied the family as a social institution, examining the roles of family structure and processes in the lives of children and elaborating on the patterns identified by social mobility researchers. The significant rise of divorce and non-marital fertility in the United States since about 1970 brought increasing interest in the impact of family transitions (e.g., divorces) and structure (e.g., presence of step-siblings) on the lives of youth. The status attainment tradition included measurements of family structure in their analyses, but interest in social mobility and families converged in work analyzing whether family structure and transitions causally impact the educational attainment, and therefore the economic productivity, of youth. Analyses generally hypothesized that larger families and non-traditional family structures (structures other than living with both biological parents) negatively affect a

youth's educational attainment. This dissertation continues and contributes to this research tradition by examining the relationships between family structure and enrollment in and completion of 4-year college in the late 20th century.

CHAPTER 1

FAMILY SIZE AND HIGHER EDUCATIONAL ATTAINMENT

Introduction

Social scientists have long been interested in the role of the family in the status attainment processes, and have theorized and investigated the relationship between family size and children's educational attainment since the mid-20th century (Anastasi 1956; Becker, Duesenberry and Okun 1960; Becker and Lewis 1973). These scholars hypothesized a trade-off between "quality and quantity" of children, such that children from larger sibships are likely to obtain less education than children from smaller sibships. This relationship has since been examined by scholars in multiple disciplines, with conclusions contingent on setting, methodology, and the particular educational outcome investigated (Jaeger 2008; Steelman et al. 2002).

The impact of family size on life outcomes such as education remains of significant interest in post-industrial economies such as the United States. The United States values equality of opportunity for its youth and demonstrates this value through its educational policies. Beginning with the post-World War II G.I. Bill, the U.S. federal government has spent significant amounts of taxpayer money on policies encouraging educational attainment (Goldin and Katz 2008). As the monetary returns to post-secondary education have increased over the 20th century, policy makers have increasingly focused on the importance of equality of opportunity to achieve higher education (Card 1999).

In particular, programs such as Pell Grants, Title IX, and subsidized federal loans for education were designed with the explicit goal of equalizing opportunity for educational achievement across races/ethnicities, genders, and incomes. Social scientists and policy-makers have spent significant amounts of time researching the

effects of these policies, but their efforts to measure program effectiveness have been complicated by the rapidly changing cost structure of higher education and differential use of the programs among demographic and socioeconomic groups (Dynarski and Scott-Clayton 2006). The hierarchical nature of education has also muddied the interpretation of these findings—no one may attend a four-year college without graduating from high school, but these federal programs have no direct leverage to increase the pool of youth with high school diplomas (who could then access the programs).

The family size gradient in educational attainment is one indicator of the success of these policies because of the economic relationship between income, monetary expenditures per child, and number of children. If income does not change but a child is added to the household, the amount of money available to be spent on each child necessarily decreases (the amount of income is "diluted"). Therefore, larger families will have less money to spend on child-rearing and post-secondary education for each child. If we observe that children from larger families obtain less post-secondary education than children from smaller families among families with the same income, this is evidence that these U.S. policies are not completely effective in their goal of relieving cost constraints on higher education.

These policies, however, contain an array of components: each family must complete the "Free Application for Federal Student Aid (FAFSA)" for each child every year in which the child is enrolled in post-secondary education. The amount a family is expected to contribute to a child's higher education costs is a complex function of total household income, allowances against income, and assets. The federal income allowance is comprised of taxes paid, an employment allowance, and an "income protection allowance" of approximately $4,500 per household member in 2011 dollars (US-Government 2013b). This allowance indicates that federal policy-

makers recognize the basic economic relationship between family size and money available for higher education. The question, therefore, is whether this discount has been enough to effectively enable youth from larger families to obtain post-secondary education.

Recognizing the changing demographic and educational landscape in the United States over the late 20th century, this paper investigates how the family size gradient in post-secondary educational attainment has changed between the mid-1980s and the mid-2000s using two rich nationally representative surveys. After documenting a negative relationship between family size and educational outcomes on average in both panels, the paper examines how this relationship varies by income level. The paper concludes with a discussion of the meaning of these findings with respect to cost constraints on higher education attainment.

Background: Theory

Hypotheses about the effects of family size began with Dumont's "law of capillary action" (1890), which suggested that siblings decrease the likelihood of social mobility (Bras, Kok and Mandemakers 2010). Since then, three claims about how sibship size affects educational ability or attainment have been developed: the confluence explanation, the resource dilution explanation, and the claim that the relationship is spurious (Steelman et al. 2002).

The confluence explanation, proposed by Zajonc and Markus in the 1975, hinges on the intellectual environment in which each child is raised (Zajonc 1976; Zajonc and Markus 1975). It argues that the intellectual development of children is a function of the average intelligence in the family/household. Therefore, each additional child in the family brings down the average, and families with more

children end up providing a less enriching environment for their children. This explanation also implies a birth-order effect: the first child begins his/her life in the intellectual environment supplied completely by the parents; the environments for subsequent children are polluted by the low levels of intelligence of their siblings (where intelligence is absolute, not age-normed). The confluence explanation also implies that children with no siblings should out-perform children with any siblings, but this is not empirically observed (Steelman et al. 2002). Additionally, the explanation cannot explain why there would be a sibship size effect beyond that of intellectual ability (Downey 2001).

The resource dilution explanation holds that the resources parents invest in their children are diluted by each additional child (Blake 1981). The most obvious resources are economic: for a fixed income, each additional child makes fewer dollars available for other children. These resources may affect children's educational attainment in a number of ways: via parents' ability to provide an enriching home environment (e.g., books, a computer), via parents' ability to provide educational experiences outside the household (such as trips to museums), ability to afford private school, or ability to provide resources for college attendance. A previously unrecognized possibility is that children from larger sibships may attend worse schools than they otherwise would have if they had fewer siblings. The logic behind this claim is that larger families require larger houses, which are more expensive than smaller houses. Because housing prices are related to school quality, a family could either afford a smaller house in a better neighborhood or a larger house in a worse neighborhood. The resource dilution explanation also implies that sibship density, or the closeness of sibling spacing, will be negatively related to academic achievement because resources will be reduced at a greater rate than if siblings were spaced further apart (Powell and Steelman 1990). However, there may be ceiling effects on the

sibship size relationship: once families have enough money to adequately provide educational resources, the dilution may not actually affect the academic achievement of the children. The resource dilution explanation also implies birth order effects: the average amount of resources per sibling is highest for the first born, and decreases with every child thereafter.

Family size may play an important role in educational attainment beyond its economic effects, however. After income, parental attention is the most commonly cited resource that additional children will dilute (Blake 1985). Because parental attention is difficult (and costly) to observe, significantly less research has empirically investigated the effects of sibship size on attention. However, scholars agree that the amount of direct parent-child interaction (especially verbal interaction and reading time) has significant effects on children's intellectual development and school readiness (Brooks-Gunn and Markman 2005).

It seems likely that the direct effects of parental attention on educational outcomes are concentrated in the early years of children's lives. After all, it is during these years that children's minds are changing most rapidly and they are developing many of the cognitive skills that they will rely on later in life. There has recently been increasing attention paid to the importance of development in early life (Brooks-Gunn and Duncan 1997; Cunha and Heckman 2007). However, there may be significant indirect effects of parental attention via the likelihood that youth will engage in delinquency or otherwise endanger their educational future in the high school years. Older children in large sibships may be especially at risk because parents need to pay closer attention to younger siblings once older children are moderately self-sufficient. Additionally, older children (especially females) may be called upon to participate in child care for younger siblings, taking away time that might otherwise be spent studying or engaged in education-related activities.

These explanations lead to a number of predictions for the empirical analysis to follow. First, the most basic prediction: sibship size will be negatively related to 4-year college enrollment and completion, net of demographic factors, income, and academic/intellectual ability. Second, because the effect of sibship size may be moderated by academic/intellectual ability, the sibship size effect will decrease once these factors have been adjusted for in the model. Third, there will be a significant interaction between sibship size and sibship density, such that the sibship size effect is stronger among closely spaced sibships. Fourth, if there are income ceiling effects on the sibship size effect, there will be a significant interaction between sibship size and income, such that the sibship size effect is weaker among those with higher incomes. Finally, birth order will be negatively related to 4-year college enrollment and completion.

Background: Evidence

The potential importance of family size for educational and occupational attainment has been recognized in a number of academic contexts. It was often included in early status attainment models, retaining a negative and significant relationship to occupational status once other demographic factors had been taken into account (Blau and Duncan 1967; Featherman and Hauser 1978). However, it was generally treated as a control variable in these works, and not as having particular interest in and of itself (Downey 1995). Although theories of why the number of siblings in a family had a causal effect on intelligence and/or educational attainment were proposed relatively early (see below), the preponderance of evidence for the relationship did not accrue until the 1980s with the work of Blake, along with that of Steelman and colleagues (Blake 1981; Blake 1985; Mercy and Steelman 1982; Steelman 1985; Steelman and Mercy 1983). Blake's 1989 book served as the capstone

to this line of research, analyzing multiple national datasets and documenting a consistent and strong negative relationship between sibship size and achievement in school (Blake 1989). While researchers recognized that the broad empirical evidence for a negative effect of sibship size was strong, the direct evidence for the mechanisms through which the effect operated had not been firmly established.

Data and Methods

This examination of the changing relationship between family size and educational attainment analyzes data from two panels of a large, nationally-representative survey conducted by the U.S. Bureau of Labor Statistics: the National Longitudinal Survey of Youth 1979 (NLSY79) and the National Longitudinal Survey of Youth 1997 (NLSY97). These surveys contain extensive and detailed data about each respondent's family, living situation, and educational experiences.

The NLSY79 was a probability sample of youth ages 14-22 in 1979, with additional surveys annually until 1994 and biennially since then. The full sample is comprised of four sub-samples, cross-sectional, ethnic, military, and poverty. Following NLSY documentation guidelines, to produce analyses comparable to the NLSY97 data, this paper uses only the cross-sectional and ethnic sub-samples of the NLSY79 (National Longitudinal Surveys 2013b). The analytic sample has been restricted to respondents aged 14-18 in 1979 and has been truncated at year 1991 to contain data on the same age ranges as the later NLSY97 data (described below). Only respondents remaining in the sample after 13 waves are included in the analysis, yielding a final sample size of N= 5,718. These sample restrictions ensure that any observed changes between the studies are not artifacts of differences in cohort age or number of survey waves.

The NLSY97 was a probability sample of youth ages 12-16 in 1997, with additional surveys annually since then. The full sample is comprised of two sub-samples, cross-sectional and ethnic, both of which will be used for the analysis. As with the 1979 panel, only respondents remaining in the sample after 13 waves are included in the analysis, yielding a final sample size of N=5,364.

The primary variable of interest for the analysis is "sibship size" – the total number of siblings drawing from the family's resource pool, including the respondent. Sibship size data were collected differently in the two panels of the survey. In the NLSY79, each respondent was asked directly the number of siblings he/she had: "How many (living) brothers and sisters do you have?" If the respondent was unsure about whom to include in this measure, they were instructed to "think of whomever you consider as your brothers and sisters." Because a significant majority (73 percent) of households had a traditional two parent structure, the reported number reflects full biological and, presumably, adopted siblings for those respondents. We are unable to determine whether step-siblings were included in the responses because the only other source of data is a roster of those living in the respondent's household at the time of the survey, so any older siblings who were not living in the household are unobserved. In the NLSY97 panel, data were collected on every sibling relationship type for siblings living in or out of the respondent's household. Because we are unable to determine precisely who was included in the early panel's sibling count, making an equivalent variable in the later panel is not possible. Therefore, we use a measure that is likely to reflect the set of siblings sharing the respondent's economic resource pool for college costs – full biological siblings, half siblings who have the same father, and adopted siblings. Step-siblings were not included in the count because it seems unlikely that a step-parent would contribute significantly to a step-child's college costs. If any parent assisted the youth with college costs, it seems more likely that

biological parents would even though they were no longer married. The findings for the later panel are robust to the use of alternative measures of sibling size, including all siblings of any type or only full biological siblings. Sibship sizes decreased markedly between 1979 and 1997, consistent with literature on changes in total fertility in the United States over the 20th century (Wetzel 1990). In the 1979 panel, 30% of respondents (an estimated 21% of youth aged 14-18 in 1979) had more than 5 siblings present in the household. In the 1997 panel, 3.21% of respondents (an estimated 2.5% of youth aged 12-16) had more than 5 siblings present in the household. In the 1997 panel, youth from the largest families have a low probability of obtaining higher education based on observable characteristics other than family size – the sibship size variable was top-coded at 5 to ensure that any observed sibship size effects are not being disproportionately driven by these largest family youths' particularly low *a priori* probability of obtaining higher education. The sibship size variable was also top-coded at 5 in the earlier panel to maintain the comparability of the analyses. The top-coding restricts the variance of the sibship size, so the model findings for sibship size are necessarily conservative – the regression slopes will be uniformly lower than they would without this adjustment.

This analysis also includes two other aspects of sibling composition that past research indicates may be relevant for the relationship between sibship size and educational attainment (de Haan 2010; Powell and Steelman 1990). The first aspect is whether or not siblings are closely spaced. The resource dilution explanation for the sibship size penalty posits that for a family's fixed amount of income and parental availability, additional children dilute the amount available for each child. If costs and need for attention vary with time, closely spaced children will require more resources at the same time than would less closely spaced children. In the context of higher education, the primary impact of closely spaced children would occur when more than

one child is in college at a time, thus requiring more resources than if only one child was in college. Therefore, this analysis operationalizes "closely spaced" as having another sibling who was born within 3 years of a given respondent because that is the maximum spacing in which children could simultaneously attend college. The second aspect of sibling composition that may be relevant to the relationship between sibship size and educational attainment is birth order. There are a number of reasons birth order could impact educational attainment, and many have been proposed in the literature (Steelman et al. 2002). In the context of higher education, the most likely mechanism through which birth order would affect educational attainment may be via resource depletion over time. For example, a family's savings to pay for college may be depleted by older siblings, leaving younger siblings with greater financial burden and potentially forcing them to absorb more college cost via loans. Operationalizing birth order is difficult, however, because it is highly collinear with sibship size (e.g., one cannot have a birth order of "third" without having a sibship size of 3). Therefore, this analysis adopts the approach of Booth and Kee, who create a birth order index that normalizes birth order by the number of siblings present in the household (Booth and Kee 2008). Specifically, the index value for a given respondent is calculated as his birth order divided by the average birth order in the family. This means that higher values for the index are associated with younger age because someone born third is younger than someone born second. Because this analysis top-codes sibship size at 5, the index ranges from 1/3 to 1-2/3 in the models below.

The primary outcome variables of interest for this analysis are 4-year college enrollment and 4-year college completion. The paper does not analyze 2-year college attendance due to (1) the historical U.S. policy focus on 4-year college attendance over 2-year college attendance, (2) the large increase in 2-year college attendance over the time periods, (3) the relatively low cost of 2-year colleges relative to 4-year colleges

(which implies that the likelihood of cost constraints on attendance are lower), and (4) the additional analytical complexity added by analyzing which youth attend which type of college. Of course, attending any form of higher education is encouraged by the U.S. government, whose programs serve both college types. However, the greatest returns to education are concentrated among 4-year college degrees, which are currently of primary importance with respect to educational inequality in the U.S. population.

One particularly valuable item that was collected in both panels is respondents' scores on the Armed Services Vocational Aptitude Battery (ASVAB) test. Direct evidence of intellectual ability is rare in surveys of this size and representativeness. The arithmetic reasoning scale score from this test is included in some of the estimated models below to adjust for differences in intellectual ability. Following NLSY documentation recommendations, the scaled versions of the score are used, which allows comparisons across individuals (National Longitudinal Surveys 2013a).

All analyses use panel sample weights to make the data representative of the population from which the survey was drawn. Only respondents remaining in the sample after 13 waves were included in the estimates, but the weights are designed to maintain the sample's representativeness.

All item-missing data for model variables were singly imputed (see Appendix for details). For continuous variables, a regression model on the outcome was estimated among respondents with data, and the values for those missing data were imputed based on predicted values from the regression model. For binary variables, a logistic regression model with the variable as the outcome was estimated among those with missing data, and the values for respondents missing the variable were imputed based on the predicted probabilities from the model. For categorical variables, a hot-deck procedure was used to randomly select a variable value from sample members

who matched the individual with missing data on a number of characteristics. The characteristics were race/ethnicity, whether their residence was in a rural area, whether they attended public or private school, whether they graduated from high school, whether they enrolled in a 4-year college, and whether they graduated from a 4-year college.

After presenting descriptive statistics and bivariate trends, the general strategy for the analysis is to estimate identical regression models on each panel and examine the marginal slopes of the sibship size variable overall and by income subpopulations. The focus of the analysis will be changes in the patterning of and marginal relationship between sibship size and educational attainment over the 18 years separating the survey cohorts.

Descriptive Statistics

Tables 1 and 2 display summary statistics for the two survey panels. The changes in population composition are consistent with documented trends – the U.S. has become more racially/ethnically diverse, with a significant increase in Hispanic residents. Families have fewer children, and fewer families have a traditional two-biological-parent structure. There are more immigrants present in the later panel and women are having their first children at later ages. Schools are smaller in the later panel, with lower student-teacher ratios, and more students are retained in grade. Non-GED high school completion has increased to a little over 80%, consistent with the findings of Heckman and LaFontaine (2010). Enrollment and completion at 4-year colleges have both increased.

Table 1 - Summary Statistics - National Longitudinal Surveys of Youth

		1979		1997	
Variable	**Values**	**Proportion**	**SE**	**Proportion**	**SE**
Race/Ethnicity	White/Other	0.79	0.006	0.72	0.007
	Black	0.14	0.005	0.15	0.005
	Hispanic	0.07	0.003	0.13	0.005
Gender	Male	0.51	0.008	0.51	0.008
	Female	0.49	0.008	0.49	0.008
Family Structure	Traditional	0.73	0.008	0.55	0.009
	Step-parent	0.09	0.005	0.13	0.006
	Single Mom	0.13	0.006	0.23	0.007
	Other	0.05	0.004	0.08	0.005
Number siblings of respondent	0	0.03	0.003	0.16	0.006
	1	0.16	0.007	0.37	0.008
	2	0.24	0.008	0.27	0.008
	3	0.21	0.008	0.12	0.006
	4	0.36	0.009	0.08	0.005
Immigrant Generation	1	0.04	0.003	0.05	0.003
	2	0.05	0.004	0.10	0.005
	3+	0.90	0.005	0.85	0.006
Region	Northeast	0.20	0.008	0.18	0.007
	North Central	0.31	0.009	0.28	0.008
	South	0.32	0.009	0.34	0.008
	West	0.17	0.007	0.20	0.007
Rural	Yes	0.22	0.008	0.28	0.008
	No	0.78	0.008	0.72	0.008
School type	Private/Parochial	0.06	0.004	0.10	0.005
	Public	0.94	0.004	0.90	0.005
Student-teacher ratio	<14	0.07	0.005	0.23	0.007
	14 to <18	0.24	0.008	0.40	0.008
	18 to <22	0.51	0.009	0.22	0.007
	22+	0.18	0.007	0.14	0.006
School size	<100	0.01	0.001	0.01	0.001
	100-299	0.04	0.004	0.07	0.004
	300-499	0.05	0.004	0.13	0.006
	500-749	0.12	0.006	0.24	0.007
	750 to 999	0.14	0.006	0.18	0.006
	1000+	0.65	0.009	0.38	0.008
Retained in grade	No	0.95	0.003	0.86	0.005
	Yes	0.05	0.003	0.14	0.005

Table 2 - Summary Statistics - National Longitudinal Surveys of Youth

	1979		1997	
Variable	**Mean**	**SE**	**Mean**	**SE**
Graduate HS	0.76	0.007	0.83	0.006
Enroll 4-year College	0.40	0.008	0.48	0.008
Complete 4-year College	0.24	0.007	0.32	0.008
Age	16.25	0.020	14.68	0.021
Mother's age at R's birth	26.33	0.107	25.86	0.086
Mother's age at first birth	21.58	0.074	23.35	0.078
Mother's Highest Grade	11.52	0.049	13.12	0.046
Annual income (thousands of 1990 dollars)	36.65	0.447	42.19	0.596
ASVAB Math Scale Score	-101.81	16.986	-258.19	14.851

Table 3 displays mean sibship sizes at the household level across the two survey panels.[1] The overall mean sibship size in the population has markedly decreased, from 3.57 in 1979 to 2.46 in 1997. This finding coincides with research on changing U.S. family sizes over the 20th century (Wetzel 1990). In the earlier panel, white and other race population members have families that are approximately .55 children smaller than those of black and Hispanic population members. This pattern of smaller families among white population member remains the same, albeit compressed, in the later panel. However, mean family sizes for black population members fell further than those for Hispanic population members, with the latter having an average of approximately .25 more children than black population members in the later panel.

Table 3 also demonstrates that mean sibship sizes decrease as income increases. However, the compression of the sibship size distribution in the later panel means the relationship is much weaker – in the early panel, population members in the lowest income quartile had about .54 larger sibship sizes than did those in the highest

[1] This and all subsequent analyses of sibship size are calculated using sibship size top-coded at 5, as described in the previous section.

income quartile. In the later panel, the income differential in sibship sizes between bottom and top income quartiles decreased to approximately .17.

Table 3 - Mean Sibship Size by Demographics – National Longitudinal Surveys of Youth

		1979		1997	
Variable	**Values**	**Mean**	**SE**	**Mean**	**SE**
Overall		3.57	0.023	2.46	0.019
Race	White/Other	3.46	0.027	2.40	0.023
	Black	4.00	0.038	2.52	0.044
	Hispanic	3.99	0.045	2.77	0.050
Income Quartile	1 Lowest	3.95	0.049	2.55	0.047
	2	3.71	0.048	2.47	0.042
	3	3.42	0.044	2.50	0.037
	4 Highest	3.41	0.039	2.38	0.031

Table 4 displays the college outcome variables of interest by race, income quartile, and number of siblings. As noted earlier, enrollment in 4-year colleges has increased among the population between the first and second panels. However, these gains have been larger among white/other race and Hispanic population members, with rates for black population members increasing by about .04. Completion of 4-year colleges has increased the same amount over time as enrollment, with gains for all three race/ethnicity groups. Again, however, the gains were least for black population members (an increase of .04 as compared to about .08 for the other race/ethnicity groups).

As recently documented, there is a steep and increasing gradient in college enrollment and completion by income (Bailey and Dynarski 2011; Reardon 2011). For enrollment, there is a difference of .41 between the highest and lowest income quartiles in the early panel, as compared to a difference of .53 in the later panel. Similarly, for completion there is a difference of .37 between the highest and lowest income quartiles in the early panel, as compared to a difference of .48 in the later

panel. The concentration of educational gains among the upper half of the income distribution is apparent, with gains of about .13 in enrollment (versus .04 in the lower half) and .13 in completion (versus .03 in the lower half).

Table 4 – 4-Year College Enrollment and Completion by Demographics – National Longitudinal Surveys of Youth

		4-year Enrollment				4-year Completion			
		1979		1997		1979		1997	
Variable	**Values**	**Proportion**	**SE**	**Proportion**	**SE**	**Proportion**	**SE**	**Proportion**	**SE**
Race	White/Other	0.43	0.009	0.53	0.010	0.27	0.008	0.38	0.010
	Black	0.32	0.012	0.36	0.015	0.13	0.009	0.17	0.012
	Hispanic	0.26	0.014	0.34	0.017	0.10	0.010	0.18	0.014
Income Quartile	1 Lowest	0.20	0.014	0.21	0.013	0.07	0.009	0.09	0.010
	2	0.28	0.014	0.35	0.016	0.14	0.011	0.19	0.013
	3	0.34	0.014	0.47	0.016	0.17	0.011	0.30	0.015
	4 Highest	0.61	0.014	0.74	0.013	0.44	0.014	0.57	0.015
Number siblings	0	0.48	0.044	0.42	0.019	0.27	0.039	0.26	0.017
	1	0.53	0.020	0.55	0.013	0.34	0.019	0.38	0.013
	2	0.46	0.016	0.48	0.016	0.30	0.015	0.35	0.015
	3	0.39	0.017	0.44	0.024	0.24	0.015	0.26	0.022
	4+	0.30	0.011	0.33	0.027	0.15	0.009	0.21	0.025

Table 4 replicates a pattern described in past literature on the relationship between sibship size and educational attainment – instead of a monotonic negative relationship between the two, we observe that youth from single child families are less likely to enroll or complete than are youth from two-child families (Steelman et al. 2002). However, there is a net negative relationship between sibship size and 4-year enrollment in the early panel, where 48 percent of children from the smallest families enroll, as compared to 30 percent of children from the largest families. A smaller enrollment gradient remains in the later panel, with lower enrollment among the smallest families and increases among the largest. There is also a net negative relationship between sibship size and college completion. Examining the temporal changes, we observe an opposite pattern than that for 4-year college enrollment – the

gains are concentrated among the largest families, with smaller families showing declines.

Overall, Table 4 paints a detailed demographic picture of 4-year college attendance and completion between the early 1980s and 2000s. Youth have responded to the increasing cultural and monetary value of 4-year college degrees by increasing attendance and completion. However, these gains have been unevenly distributed among the population, primarily accruing to those with higher incomes and population members who are white/other race. As documented in the literature, some of these changes have been driven by increased educational attainment among women (Goldin and Katz 2008). Indeed, women now outpace men in completed education in the United States. However, unless families disproportionately direct college expenditures toward female children, cost constraints on higher educational attainment affect youth of both genders. If youth from larger families are more likely to forgo college due to cost concerns, this remains an important source of inequality of educational attainment and a drag on the development of the United States workforce. In the next section, we turn to analysis of the relationship between sibship size and educational attainment.

Regression Analysis - Enrollment

The descriptive results presented in the last section portray a net negative relationship between sibship size and educational attainment. This pattern has been extensively documented and is not surprising; however, as described above, whether this relationship is causal at a fundamental level remains to be settled in the extant literature (Guo and VanWey 1999).

The regression analyses below do not represent clean causal estimates. While they include an extremely rich set of demographic, family, and individual adjustment variables, their estimates are biased if important unobserved factors play an important

role. The most commonly cited omitted variables in such models are intellectual ability and non-cognitive traits such as persistence and ability to concentrate (Cunha and Heckman 2007). In particular, if youth with low intellectual ability or poor cognitive traits are disproportionately likely to come from large sibship size households before the putatively causal effects of sibship size have acted, then the models will erroneously ascribe the relationship due to ability/non-cognitive traits to sibship size.

However, the final models adjust for mathematical ability at the beginning of the survey panel by including each respondent's ASVAB math scale score and whether the respondent was ever retained in grade. Because these variables were measured in the respondents' early teens, any causal effect of sibship size that operates via decreased academic ability will be attributed to these variables instead of the sibship size variable. This means that the sibship size effect estimate is doubly conservative – the sibship size variable has been top-coded at 5 to prevent results being driven by the largest families, and the sibship size effect that operates via academic ability will not be attributed to sibship size. These precautions mean that the analyses below may be interpreted as conservative estimates of the relationship between sibship size and higher educational attainment.

Table 5 displays six logistic regression models with estimated odds ratios, where each model adds adjustment variables to the previous. The first model replicates the negative bivariate relationship between sibship size and enrollment in 4-year colleges presented in Table 4, indicating that the log odds of enrolling decrease by a factor of about .75 for each additional sibling one has. The second model adds the other two measures of sibling configuration, density and the birth order index described above. Sibship size remains significantly negative, and having a sibling within 3 years of oneself is negative and significant.

Table 5 - 4-Year College Enrollment Regression Models - National Longitudinal Survey of Youth 1979

Variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Sibship Size	0.745***	0.769***	0.794***	0.824**	0.895	0.900
	(0.0224)	(0.0260)	(0.0426)	(0.0656)	(0.0775)	(0.0839)
Sibling within +/- 3 years		0.825**	0.957	0.637*	0.603*	0.559**
		(0.0714)	(0.223)	(0.155)	(0.158)	(0.156)
Sibling order index		0.906	0.904	0.819**	0.558***	0.678***
		(0.0745)	(0.0744)	(0.0699)	(0.0684)	(0.0889)
Sibship size X Sibling w/in 3 years			0.955	1.030	1.075	1.099
			(0.0625)	(0.0711)	(0.0799)	(0.0876)
Income ($1k of 1990 dollars)				1.035***	1.031***	1.028***
				(0.00694)	(0.00693)	(0.00731)
Sibship Size X Income				1.000	0.998	0.997
				(0.00174)	(0.00165)	(0.00176)
+Demographics & School Vars					√	√
+ASVAB & Retained in Grade						√
N	5718	5718	5718	5718	5718	5718
Pseudo-R2	0.0220	0.0231	0.0232	0.102	0.190	0.278

* p<0.05; ** p<0.01; *** p<0.001

The third model adds the predicted interaction between sibship size and sibling density, which is not significant, indicating that the observed sibship size relationship is present for both closely and widely spaced sibships. The fourth model adds the predicted interaction between income and sibling size; income is significant and positive, but the interaction is not significant. This means that the observed sibship relationship is approximately equal for youth of all incomes. The main effect of having a closely spaced sibling remains negative and significant. Once the interactions have been included, the birth order index is significant and negative as predicted, indicating that younger siblings are less likely to enroll in 4-year college than older siblings.

The fifth model adds the demographic and school adjustment variables from Tables 1 and 2. Once these adjustments are made, the main effect of sibship size among widely spaced siblings is no longer significant. The main effect of having a closely spaced sibling remains significant and negative, as does the birth order index. The sixth model adds the ASVAB mathematics score and whether the respondent was ever retained in grade. These additions significantly boost the explanatory power of the model, and the patterns observed in the fifth model remain.

Table 6 displays the same set of six logistic regression models as Table 5, but for the 1997 panel. The first model indicates that the bivariate relationship between sibship size and 4-year college enrollment is weaker in the 1997 panel than in the earlier panel, as described in Table 4. The second model adds the other components of

Table 6 - 4-Year College Enrollment Regression Models - National Longitudinal Survey of Youth 1997

Variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Sibship Size	0.894***	0.860***	0.967	0.843**	0.850**	0.882
	(0.0252)	(0.0269)	(0.0411)	(0.0648)	(0.0675)	(0.0695)
Sibling within +/- 3 years		1.234***	2.095***	1.229	0.998	0.950
		(0.0882)	(0.360)	(0.227)	(0.202)	(0.204)
Sibling order index		0.894	0.885	0.781***	0.679***	0.842
		(0.0759)	(0.0753)	(0.0726)	(0.0879)	(0.117)
Sibship size X Sibling w/in 3 years			0.806***	0.980	1.042	1.020
			(0.0495)	(0.0657)	(0.0772)	(0.0796)
Income ($1k of 1990 dollars)				1.026***	1.016***	1.014***
				(0.00524)	(0.00496)	(0.00439)
Sibship Size X Income				1.003	1.001	1.000
				(0.00184)	(0.00161)	(0.00141)
+Demographics & School Vars					√	√
+ASVAB & Retained in Grade						√
N	5364	5364	5364	5364	5364	5364
Pseudo-R2	0.00291	0.00485	0.00699	0.118	0.198	0.309

* p<0.05; ** p<0.01; *** p<0.001

sibling configuration, with birth order significantly negative but having a closely spaced sibling unexpectedly positively associated with being more likely to enroll in college. This could reflect a tendency for families with large a priori probabilities of having children who attend college choosing to have exactly two closely spaced children; Table 1 indicates that having one sibling is the modal sibship size for youth in this panel.

The third model adds the interaction between sibship size and having a closely spaced sibling, which is significant negative. Because the main coefficient of sibship size is not significant, this means that the sibship size relationship only occurs among closely spaced siblings when demographic factors are not adjusted for. The main effect of having a closely spaced sibling remains significant and positive. The fourth model adds income and the interaction between income and sibship size, which is not significant. The birth order index is significant and negative, as expected. The positive relationship between having a closely spaced sibling and enrollment in a 4-year college is now insignificant, suggesting that the relationship is explained away by income. Because income is positively related to enrollment in a 4-year college, this lends modest support to the earlier speculation that families with large a priori probabilities of having college-enrolling children are more likely to have closely spaced children.

The fifth model adds the demographic and school variables. The patterns remain the same as in the fourth model. The negative relationship between sibship size and enrollment in 4-year college remains significant once the demographic and school variables have been added to the model. This provides evidence that the relationship is not spurious. Higher birth order respondents are less likely to enroll even when they have similar demographic and school characteristics as lower birth order respondents, which could be due to multiple factors. It could indicate that parental resources for

college have been depleted after helping pay for older siblings to attend college, or it could be due to younger children observing their siblings struggle to succeed and change their beliefs about their own probability of successfully completing college. The latter possibility is supported by the fact that a significant minority of youth who enter college do not graduate, with most of them taking on debt to do so (Adelman 2006).

Once the models include ASVAB mathematics score and whether the respondent has been retained in grade, the sibship size coefficient becomes insignificant for widely spaced sibships, suggesting that the mechanism of action for sibship size may be decreased academic ability. Specifically, many more respondents are retained in grade in the later panel, and individuals from larger sibships may have more problems succeeding in academic settings, which would lead them to be retained in grade and subsequently less likely to attend 4-year college. Additionally, birth order becomes insignificant when the academic ability variables are included in the model, suggesting that birth order effects may also act via academic ability.

Table 7 displays the marginal slope for the sibship size variables across the six logistic regression models for both panels. These values reflect the instantaneous rate of change in probability associated with a small increase in sibship size (as compared to odds ratios in the logistic models above), averaged over the characteristics of the sample. The sibship coefficients in the models above reflect the sibship relationship conditional on the model variable values (i.e., holding the variable values constant), whereas the values in Table 7 reflect the change in probability associated with a change in sibship calculated for each sample member, then averaged. Mirroring the models above, there is a highly-significant negative relationship between sibship size and 4-year college enrollment for both panels (in the first row of the table). However,

the size of the effect is a little less than half as large in the later panel, reflecting the compression of the sibship size distribution over time.

Table 7 - Sibship Marginal Slopes for 4-Year College Enrollment Regression Models - National Longitudinal Surveys of Youth

	1979			1997		
Model	**Slope**	**SE**	**Sig.**	**Slope**	**SE**	**Sig.**
Sibship	-0.069	0.007	<.001	-0.028	0.007	<.001
+Close & Order	-0.061	0.008	<.001	-0.037	0.008	<.001
+CloseXSibship	-0.062	0.008	<.001	-0.039	0.008	<.001
+IncomeXSibship	-0.039	0.007	<.001	-0.017	0.008	0.023
+Demographics/School Variabless	-0.024	0.008	0.003	-0.017	0.008	0.040
+ASVAB & Retained in grade	-0.021	0.008	0.007	-0.017	0.007	0.021

In both panels, the relationship between sibship size and 4-year college enrollment remains negative and significant in all subsequent models. This means that even when demographically similar youth who attend similar schools, have similar family incomes, and have similar academic abilities are compared, those from larger families have significantly lower probabilities of enrolling in a 4-year college. The results from the final model imply that the probability of a youth from a sibship size of 5 or more enrolling in 4-year college would be .084 smaller than that from an equivalent youth from a sibship size of 1 in the early panel. To put this difference in perspective, it is about ¾ the size of the widely studied black-white gap in probability of enrollment and about 1/5 the size of the gap in probability between the highest and lowest income quartiles. This is a sizable source of inequality in enrollment in higher education. In the later panel, youth from a sibship size of 5 or more are .068 less likely to enroll in 4-year college than youth from a sibship size of 1. This difference is approximately 1/3 the size of the black-white difference in probability of enrollment and about 1/8 the size of the gap between the lowest and highest income quartiles. The smaller magnitude in the later panel may be a result of increased access to convenient

4-year colleges and to loans for funding higher education, along with the general U.S. trend of increased college-going due to increasing returns to education.

Table 8 - Sibship Marginal Slopes by Household Income Quartile for 4-Year College Enrollment Regression Models – National Longitudinal Surveys of Youth

		1979			1997		
Model	**Income Quartile**	**Slope**	**SE**	**Sig.**	**Slope**	**SE**	**Sig.**
Sibship	1 Lowest	-0.067	0.006	0.000	-0.028	0.007	0.000
	2	-0.068	0.006	0.000	-0.028	0.007	0.000
	3	-0.069	0.007	0.000	-0.028	0.007	0.000
	4 Highest	-0.069	0.007	0.000	-0.028	0.007	0.000
+Close & Order	1 Lowest	-0.060	0.007	0.000	-0.037	0.008	0.000
	2	-0.061	0.007	0.000	-0.037	0.008	0.000
	3	-0.062	0.008	0.000	-0.037	0.008	0.000
	4 Highest	-0.062	0.008	0.000	-0.037	0.008	0.000
+CloseXSibship	1 Lowest	-0.060	0.007	0.000	-0.038	0.008	0.000
	2	-0.061	0.008	0.000	-0.039	0.008	0.000
	3	-0.062	0.008	0.000	-0.038	0.008	0.000
	4 Highest	-0.063	0.008	0.000	-0.039	0.008	0.000
+IncomeXSibship	1 Lowest	-0.027	0.008	0.001	-0.029	0.010	0.002
	2	-0.035	0.008	0.000	-0.027	0.008	0.001
	3	-0.043	0.008	0.000	-0.019	0.009	0.036
	4 Highest	-0.043	0.012	0.000	-0.002	0.014	0.886
+Demographics/School Vars	1 Lowest	-0.010	0.009	0.224	-0.021	0.009	0.020
	2	-0.017	0.008	0.037	-0.021	0.009	0.013
	3	-0.027	0.009	0.003	-0.020	0.010	0.045
	4 Highest	-0.033	0.011	0.003	-0.009	0.012	0.449
+ASVAB & Retained in grade	1 Lowest	-0.007	0.008	0.382	-0.015	0.007	0.045
	2	-0.014	0.008	0.084	-0.018	0.008	0.019
	3	-0.022	0.008	0.007	-0.019	0.008	0.021
	4 Highest	-0.030	0.010	0.005	-0.015	0.010	0.144

Table 8 examines the relationship between sibship size and 4-year college enrollment by income quartile. Interestingly, once demographic and school variables have been added to the models we see opposite patterns in the 1979 and 1997 panels. In the early panel, the negative sibship size relationship is unexpectedly significant in the higher income quartiles but not in the lowest. Because the probability of enrolling in college is so much greater in the higher income quartiles, this could be due to the likelihood of richer families having more than one child attend college. If the available

financial aid does not offset the increased expense of having multiple children in college simultaneously or families' savings for paying for college are depleted and there are credit constraints for youth obtaining funding, youth from larger families may be discouraged from enrolling in college. If this explanation is accurate, it implies that there should be a birth order effect such that younger siblings are less likely to enroll than older siblings, which is observed in the model from Table 5 containing all adjustment variables.

In contrast to the earlier panel, the relationship in the later panel reveals the expected interaction between income and sibship size: the sibship size slope becomes insignificant in the highest income quartile. One possibility is that this reflects lower income youth becoming discouraged after observing their older siblings have a relatively low likelihood of succeeding in college. However, this implies that the birth order negative relationship should remain significant in all models, but it is not in the final model that includes measures of academic ability. Instead, it may indicate that individuals from higher income families either have the resources to pay for college outright or are more willing to take on debt to obtain a college degree. If higher income parents are more willing to assume some of the college debt, that means higher income youth must shoulder less of it, which may make them have a higher propensity to enroll regardless of sibship size (in addition to other cultural expectations of participation in college). Additionally, in 1997 the first non-need based financial aid policies in the form of tax credits for tuition expenses were passed into law; these credits are expected to go primarily to middle and upper income households (Dynarski 2000). Even though educational grants and loans are available for low-income youth (such as Pell grants, Stafford loans, etc.), lower income youth may be more reluctant to take on debt to obtain a college degree.

In summary, regression analysis of 4-year college enrollment indicates that, on average, there is a significant negative relationship between sibship size and 4-year college enrollment that has only modestly changed between the mid-1980s and early 2000s. However, the income groups among which this negative relationship appears have reversed in that time period, switching from the highest to the lowest quartiles.

Regression Analysis - Completion

We now examine the relationship between sibship size and completion of 4-year college. Table 9 displays the odds ratios for the same series of logistic regression models that were estimated in the enrollment analysis for the early panel. As before, the first model replicates the negative bivariate relationship between sibship size and

Table 9 - 4-Year College Completion Regression Models - National Longitudinal Survey of Youth 1979

Variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Sibship Size	0.732***	0.742***	0.788***	0.792**	0.885	0.875
	(0.0247)	(0.0287)	(0.0445)	(0.0723)	(0.0888)	(0.0962)
Sibling within +/- 3 years		0.922	1.207	0.747	0.675	0.637
		(0.0919)	(0.301)	(0.202)	(0.202)	(0.202)
Sibling order index		0.963	0.959	0.832*	0.549***	0.700**
		(0.0949)	(0.0944)	(0.0866)	(0.0849)	(0.115)
Sibship size X Sibling w/in 3 years			0.916	0.999	1.065	1.087
			(0.0653)	(0.0781)	(0.0918)	(0.103)
Income ($1k of 1990 dollars)				1.038***	1.033***	1.029***
				(0.00786)	(0.00744)	(0.00763)
Sibship Size X Income				1.000	0.998	0.998
				(0.00195)	(0.00178)	(0.00183)
+Demographics & School Vars					√	√
+ASVAB & Retained in Grade						√
N	5718	5718	5718	5718	5718	5718
Pseudo-R2	0.0236	0.0238	0.0241	0.134	0.238	0.340

* p<0.05; ** p<0.01; *** p<0.001

college completion displayed in Table 4. The second model adds the other sibling configuration variables, having a closely spaced sibling and birth order, neither of which are significant. The third model adds the predicted interaction between sibship size and having a closely spaced sibling, which is not significant. The fourth model adds income and the predicted interaction between income and sibship size; the income coefficient is significant and positive, but the interaction is not significant.

The main effect of sibship size among widely spaced sibships remains significant and negative until the fifth model, where demographics and school variables are added. The birth order index is significantly negative and quite strong, with a unit increase associated with a 45% reduction in the log odds of completing 4-year college. This pattern of coefficients remains once the academic ability variables have been added, with the effect of birth order remaining significant but weakening.

Table 10 displays the same series of models estimated on the 1997 panel. The first model displays a smaller negative relationship between sibship size than that found in the earlier panel. The second model replicates the finding for enrollment that closely spaced siblings are more likely to complete college, which remains in the third model. This relationship was found to be due to demographic composition; those families with a priori high likelihoods of having college-going children were more likely to have closely spaced siblings. The third model finds the expected negative interaction between sibship size and having a closely spaced sibling, meaning that the sibship size relationship is stronger among closely spaced sibships.

The fourth model adds income and the interaction between income and sibship size. Both are significant and in the expected direction, with the negative sibship size relationship weaker among those with higher incomes. The fifth model adds the demographic and school variables, which make the main effect of income and closely

spaced sibling no longer significant. However, sibship size and the interaction between sibship size and income remain significant and in the expected direction.

Table 10 - 4-Year College Completion Regression Models - National Longitudinal Survey of Youth 1997

Variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Sibship Size	0.912***	0.858***	0.985	0.745***	0.753***	0.771***
	(0.0273)	(0.0296)	(0.0446)	(0.0665)	(0.0692)	(0.0721)
Sibling within +/- 3 years		1.368***	2.484***	1.446*	1.113	1.095
		(0.108)	(0.459)	(0.299)	(0.252)	(0.259)
Sibling order index		0.968	0.957	0.859	0.819	1.031
		(0.0914)	(0.0898)	(0.0879)	(0.113)	(0.146)
Sibship size X Sibling w/in 3 years			0.780***	0.963	1.044	1.015
			(0.0523)	(0.0745)	(0.0902)	(0.0903)
Income ($1k of 1990 dollars)				1.017***	1.006	1.004
				(0.00506)	(0.00404)	(0.00403)
Sibship Size X Income				1.006***	1.004***	1.003**
				(0.00188)	(0.00147)	(0.00142)
+Demographics & School Vars					√	√
+ASVAB & Retained in Grade						√
N	5364	5364	5364	5364	5364	5364
Pseudo-R2	0.00188	0.00527	0.00789	0.128	0.222	0.319

* p<0.05; ** p<0.01; *** p<0.001

The final model adds the academic ability variables and displays the same coefficient patterns as the previous model, with the magnitudes slightly moderated.

Table 11 presents the marginal slopes for sibship size in regression models of 4-year college completion. In the 1979 panel, there is a strong negative bivariate relationship. As adjustment variables are added to the model, the relationship weakens, with a slope of -0.016 in the final model. The results from the final model imply that the probability of a youth from a sibship size of 5 completing 4-year college would be .06 smaller than that from an equivalent youth from a sibship size of 1. This

represents about 40% of the size of the black-white gap in college completion, and about 16% of the gap between the highest and lowest income quartiles in completing a 4-year college.

Table 11 - Sibship Marginal Slopes for 4-Year College Completion Regression Models – National Longitudinal Surveys of Youth

	1979			1997		
Model						
Sibship	-0.055	0.006	<.001	-0.020	0.006	0.002
+Close/Order	-0.053	0.007	<.001	-0.033	0.007	<.001
+CloseXSibship	-0.053	0.007	<.001	-0.035	0.007	<.001
+IncomeXSibship	-0.035	0.006	<.001	-0.013	0.008	0.103
+Demographics/School Vars	-0.019	0.007	0.006	-0.012	0.008	0.130
+ASVAB & Retained in grade	-0.016	0.006	0.011	-0.013	0.007	0.061

The bivariate relationship between sibship size and 4-year college completion is significantly smaller in the 1997 panel. The relationship moderates as the models add adjustment variables, remaining moderately significant at about the same magnitude as the earlier panel in the final model containing the academic ability variables. This finding indicates that, on average, youth from larger families suffer about a 1 probability point penalty per sibling for 4-year college completion. However, as with 4-year enrollment, these overall population results may mask differences in the sibship size penalty among income quartiles of the population.

Table 12 presents the marginal sibship size slope for completion of 4-year college by household income quartile. Similar to the results for college enrollment, the relationships begin to differentiate once the demographic and school adjustment variables are included. In the earlier panel the sibship size relationship is insignificant among the lowest income quartile, moderately significant among the second quartile, and larger and significant among the top two quartiles. This pattern remains in the final model that includes the academic ability variables.

Table 12 - Sibship Marginal Slopes by Household Income Quartile for 4-Year College Completion Regression Models – National Longitudinal Surveys of Youth

		1979			1997		
Model	**Income Quartile**	**Slope**	**SE**	**Sig.**	**Slope**	**SE**	**Sig.**
Sibship	1 Lowest	-0.052	0.005	<.001	-0.020	0.006	0.002
	2	-0.054	0.005	<.001	-0.020	0.006	0.002
	3	-0.056	0.006	<.001	-0.020	0.006	0.002
	4 Highest	-0.057	0.006	<.001	-0.020	0.007	0.002
+Close/Order	1 Lowest	-0.050	0.006	<.001	-0.033	0.007	<.001
	2	-0.052	0.006	<.001	-0.033	0.007	<.001
	3	-0.054	0.007	<.001	-0.033	0.007	<.001
	4 Highest	-0.054	0.007	<.001	-0.033	0.008	<.001
+Close X Sibship	1 Lowest	-0.050	0.006	<.001	-0.033	0.007	<.001
	2	-0.052	0.006	<.001	-0.035	0.007	<.001
	3	-0.054	0.007	<.001	-0.034	0.007	<.001
	4 Highest	-0.055	0.007	<.001	-0.036	0.008	<.001
+IncomeXSibship	1 Lowest	-0.017	0.005	0.001	-0.031	0.007	<.001
	2	-0.026	0.006	<.001	-0.029	0.007	<.001
	3	-0.037	0.007	<.001	-0.019	0.008	0.020
	4 Highest	-0.047	0.012	<.001	0.015	0.016	0.352
+Demographics/School Vars	1 Lowest	-0.006	0.005	0.224	-0.021	0.006	0.001
	2	-0.011	0.006	0.059	-0.022	0.007	0.002
	3	-0.019	0.007	0.010	-0.018	0.009	0.043
	4 Highest	-0.031	0.012	0.008	0.005	0.014	0.711
+ASVAB & Retained in grade	1 Lowest	-0.005	0.004	0.259	-0.017	0.005	0.001
	2	-0.009	0.005	0.087	-0.020	0.006	0.002
	3	-0.016	0.007	0.018	-0.019	0.008	0.019
	4 Highest	-0.026	0.010	0.012	-0.001	0.012	0.910

In the later panel, we observe the predicted interaction between income and the bivariate sibship size penalty in 4-year college completion: there is a negative relationship among the lowest 3 income quartiles, decreasing to a negligible relationship in the top quartile of the income distribution. These findings mirror those for college enrollment: a negative sibship size relationship for those with incomes in the top quartile for the early panel and in the lower quartiles for the later panel. We now discuss why these patterns appear.

Discussion & Conclusion

Confirming the results of decades of research, this analysis finds that, on average, youth from larger sibships are less likely to enroll and complete college. However, it finds significant variation by income and, counter intuitively and unexpectedly, that the sibship effect is only significant among the highest income quartile in the early panel.

The finding that the sibship size effect only operated among the highest income quartile in the early panel may be due to the financial aid policy environment in the early 1980s. First, financial aid in the form of direct grants was available for lower income youth but not higher income youth. Second, educational loans for middle and higher-income students were neither subsidized nor directly available from the government, which made private sources the only way for these families to obtain credit for funding their children's higher education if they did not have savings. If these families were credit constrained and were running out of savings on older children, we should observe a birth order effect, and we do. There is also evidence that middle and upper income families have students on the margin of college attendance who do not attend due to financial pressures: (Dynarski 2000) finds that the Georgia HOPE scholarship increased enrollment of students who did not qualify for federal aid by approximately 4 percent per $1,000 of aid, an increase consistent with other literature on the enrollment effects of student aid. Finally, student loan volume exploded in 1993, when unsubsidized Stafford loans were introduced and made available to students of all income levels, while federally guaranteed loans to parents ("PLUS" loans) only marginally increased in volume, which implies pent up demand among those ineligible for subsidized student loans (College Board 1998). Together, these factors offer a plausible explanation for the finding that sibship size effects only operated among families in the top income quartile in the early 1980s: these families

had students on the margin of attendance who did not attend due to lack of family resources and constraints on available credit for college tuition. This explanation implies birth order effects because families depleted savings, and those effects are observed, net of all controls including academic ability.

By the later panel, when youth were attending college in the early 2000s, access to financial aid had significantly eased. Total volume of unsubsidized Stafford loans more than doubled, primarily due to increases in the number of borrowers (as compared to the amount of each loan) (College Board 2006). Additionally, tax credits for tuition began in the late 1990s, which are mainly used by middle and upper-income families but probably did not increase enrollment – however, they may have increased completion due to decreased depletion of college savings (Long 2004). If this is the case, we should not observe a sibship order effect in the later panel, and we do not. Additionally, low income youth may be less willing to take on educational debt than higher income youth.

In summary, this paper demonstrates that there is a negative relationship between sibship size and college enrollment and completion, on average, in both the early 1980s and the early 2000s, even when a rich assortment of adjustment variables are included in the models. However, the patterning of this relationship by income changed dramatically between the time periods: in the first, sibship size effects only appeared among the highest income quartile, while they only appeared among the lowest income quartile in the latter. The explanation for this change is the vast change in higher education aid policies between the two time periods: in the former, middle- and higher-income families were relegated to private loans for funding their children's education, which left students on the margin not attending due to lack of financing. In contrast, in the later panel student loans are readily available and taken by youth of all incomes. However, lower income youth may be reluctant to take on educational debt,

which decreases their attendance. The presence of sibship size effects among the lowest income quartile provides one plausible contributor to the widely documented expanding divide between youth of low and high income in college attendance and completion.

REFERENCES

Adelman, Clifford. 2006. "The Toolbox Revisited: Paths to Degree Completion From High School Through College." *US Department of Education*.

Anastasi, Anne. 1956. "Intelligence and Family Size." *Psychological Bulletin* 53:187-209.

Bailey, Martha J., and Susan M. Dynarski. 2011. "Inequality in Post-Secondary Education." *Whither Opportunity*:117-32.

Becker, Gary S., James S. Duesenberry, and Bernard Okun. 1960. "An Economic Analysis of Fertility." Pp. 225-56 in *Demographic and Economic Change in Developed Countries*: Columbia University Press.

Becker, Gary S., and H. Gregg Lewis. 1973. "On the Interaction between Quantity and Quality of Children." *Journal of Political Economy* 82:279-87.

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(1):84-94.

—. 1989. *Family Size and Achievement*. Los Angeles, CA: University of California Press.

Blau, Peter M., and Otis Dudley Duncan. 1967. *The American Occupational Structure*. New York: Wiley.

Booth, Alison L., and Hiau Joo Kee. 2008. "Birth order matters: the effect of family size and birth order on educational attainment." *Journal of Population Economics* 22(2):367-97.

Bras, Hilde, Jan Kok, and Kees Mandemakers. 2010. "Sibship Size and Status Attainment Across Contexts: Evidence from the Netherlands, 1840-1925." *Demographic Research* 23:73-104.

Brooks-Gunn, Jeanne, and Greg J. Duncan. 1997. "The Effects of Poverty on Children." *The Future of Children* 7(2):55-71.

Brooks-Gunn, Jeanne, and Lisa Markman. 2005. "The Contribution of Parenting to Ethnic and Racial Gaps in School Readiness." *The Future of Children* 15(1):139-68.

Card, David E. 1999. "The Causal Effect of Education on Earnings." *Handbook of Labor Economics* 3:1801-63.

College Board, The. 1998. "Trends in Student Aid." Washington, D.C.

—. 2006. "Trend in Student Aid." Washington, D.C.

Cunha, Flavio, and James J. Heckman. 2007. "The Technology of Skill Formation." *American Economic Review* 97(2):31-47.

de Haan, Monique. 2010. "Birth order, family size and educational attainment." *Economics of Education Review* 29(4):576-88.

Downey, Douglas B. 1995. "When Bigger Is Not Better: Family Size, Parental Resources, and Children's Educational Performance." *American Sociological Review* 60(5):746-61.

—. 2001. "Number of Sibling and Intellectual Development: The Resource Dilution Explanation." *American Psychologist* 56:497-504.

Dumont, Arsene. 1890. *Dépopulation et civilisation: étude démographique.* Paris: Economica.

Dynarski, Susan M. 2000. "Hope for Whom? Financial Aid for the Middle Class and Its Impact on College Attendance." *National Tax Journal* 53(3):629-61.

Dynarski, Susan M., and Judith E. Scott-Clayton. 2006. "The Cost of Complexity in Federal Student Aid: Lessons from Optimal Tax Theory and Behavioral Economics." *National Tax Journal* 59(2):319-56.

Featherman, David, and Robert M. Hauser. 1978. *Opportunity and Change*. New York: Academic Press.

Goldin, Claudia Dale, and Lawrence F. Katz. 2008. *The Race Between Education and Technology*. Cambridge, Mass.: Belknap Press of Harvard University Press.

Guo, Guang, and Leah K. VanWey. 1999. "Sibship Size and Intellectual Development: Is the Relationship Causal?" *American Sociological Review* 64(2):169-87.

Heckman, James J., and Paul A. LaFontaine. 2010. "The American High School Graduation Rate: Trends and Levels." *The Review of Economics and Statistics* 99(2):244-62.

Jaeger, M. M. 2008. "Do Large Sibships Really Lead to Lower Educational Attainment?: New Evidence from Quasi-Experimental Variation in Couples' Reproductive Capacity." *Acta Sociologica* 51(3):217-35.

Long, Bridget T. 2004. "The Impact of Federal Tax Credits for Higher Education Expenses." in *College Choices: The Economics of Where to Go, When to Go, and How to Pay For It*, edited by Caroline M. Hoxby. Chicago, IL: University of Chicago Press.

Mercy, James A., and Lala Carr Steelman. 1982. "Familial Influence on the Intellectual Attainment of Children." *American Sociological Review* 47:532-42.

National Longitudinal Surveys, The. 2013a. "Aptitude, Achievement & Intelligence Scores."

—. 2013b. "Constructing Comparable Samples across the NLSY79 and NLS&97."

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

Reardon, Sean F. 2011. "The Widening Academic Achievement Gap Between the Rich and the Poor: New Evidence and Possible Explanations." *Whither Opportunity*:91-116.

Steelman, Lala Carr. 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 Carr, and James A. Mercy. 1983. "Sex Differences in the Impact of the Number of Older and Number of Younger Siblings on IQ Performance." *Social Psychology Quarterly* 46:157-62.

Steelman, Lala Carr, Brian Powell, Regina Werum, and Scott Carter. 2002. "Reconsidering the Effects of Sibling Configuration: Recent Advances and Challenges." *Annual Review of Sociology* 28(1):243-69.

US-Government. 2013. "The EFC Formula, 2013-2014." edited by Department of Education.

Wetzel, James R. 1990. "American Families: 75 Years of Change." *Monthly Labor Review* 113(3):4-13.

Zajonc, Robert B. 1976. "Family Configuration and Intelligence." *Science* 192:227-36.

Zajonc, Robert B., and Gregory B. Markus. 1975. "Birth order and Intellectual Development." *Psychological Review* 82:74-88.

CHAPTER 2

COMPETING EXPLANATIONS FOR THE RELATIONSHIP BETWEEN FAMILY SIZE AND HIGHER EDUCATIONAL ATTAINMENT

Introduction

Social scientists have spent significant effort theorizing about and examining the relationship between the number of siblings one has and educational attainment (Anastasi 1956; Becker, Duesenberry and Okun 1960; Becker and Lewis 1973; Featherman and Hauser 1978). These scholars observed a negative relationship between sibship size and intellectual development and hypothesized that children from larger sibships are likely to obtain less education than children from smaller sibships because the number of siblings one has affects the intellectual and economic resources available during childhood. This relationship has since been examined by scholars in multiple disciplines, with conclusions contingent on setting, methodology, and the particular educational outcome investigated (Jaeger 2008; Steelman et al. 2002). Some scholars maintain that the relationship is spurious, but they make this claim by applying methodological approaches with limited generalizability and not by positing a process by which such a spurious relationship would be generated (Guo and VanWey 1999). This paper proposes three aspects of families and parents that could generate a spurious relationship between sibship size and educational attainment and evaluates the likelihood that they do so with analysis of survey data containing rich information about youth and their families. It also disaggregates the relationship between sibship size and educational attainment by race/ethnicity to examine whether demographic composition or selection play a stronger role among some youth than among others.

Background: Theory

Hypotheses about the effects of family size began with Dumont's "law of capillary action" (1890), which suggested that siblings decrease the likelihood of social mobility (Bras, Kok and Mandemakers 2010). Since then, three claims about how sibship size affects educational ability or attainment have been developed: the confluence explanation, the resource dilution explanation, and the claim that the relationship is spurious (Steelman et al. 2002).

The confluence explanation, proposed by Zajonc and Markus in 1975, hinges on the intellectual environment in which each child is raised (Zajonc 1976; Zajonc and Markus 1975). It argues that the intellectual development of children is a function of the average intelligence in the family/household. Therefore, each additional child in the family brings down the average, and families with more children end up providing a less enriching environment for their children. This explanation also implies a birth-order effect: the first child begins his/her life in the intellectual environment supplied completely by the parents; the environments for subsequent children are polluted by the low levels of intelligence of their siblings (where intelligence is absolute, not age-normed). The confluence explanation also implies that children with no siblings should out-perform children with any siblings, but this is not empirically observed (Steelman et al. 2002). Additionally, the explanation cannot explain why there would be a sibship size effect beyond that of intellectual ability (Downey 2001).

The resource dilution explanation holds that the resources parents invest in their children are diluted by each additional child (Blake 1981). The most obvious resources are economic: for a fixed income, each additional child makes fewer dollars available for other children. These resources may affect children's educational attainment in a number of ways: via parents' ability to provide an enriching home environment (e.g., books, a computer), via parents' ability to provide educational

experiences outside the household (such as trips to museums), ability to afford private school, or ability to provide resources for college attendance. A previously unrecognized possibility is that children from larger sibships may attend worse schools than they otherwise would have if they had fewer siblings. The logic behind this claim is that larger families require larger houses, which are more expensive than smaller houses. Because housing prices are related to school quality, a family could either afford a smaller house in a better neighborhood or a larger house in a neighborhood with worse schools. The resource dilution explanation also implies that sibship density, or the closeness of sibling spacing, will be negatively related to academic achievement because resources will be reduced at a greater rate than if siblings were spaced further apart (Powell and Steelman 1990). However, there may be ceiling effects on the sibship size relationship: once families have enough money to adequately provide educational resources, the dilution may not actually affect the academic achievement of the children. The resource dilution explanation also implies birth order effects: the average amount of resources per sibling is highest for the first born, and decreases with every child thereafter.

Family size may play an important role in educational attainment beyond its economic effects, however. After income, parental attention is the most commonly cited resource that additional children will dilute (Blake 1985). Because parental attention is difficult (and costly) to observe, significantly less research has empirically investigated the effects of sibship size on attention. However, scholars agree that the amount of direct parent-child interaction (especially verbal interaction and reading time) has significant effects on children's intellectual development and school readiness (Brooks-Gunn and Markman 2005).

It seems likely that the direct effects of parental attention on educational outcomes are concentrated in the early years of children's lives. After all, it is during

these years that children's minds are changing most rapidly and they are developing many of the cognitive skills that they will rely on later in life. There has recently been increasing attention paid to the importance of development in early life (Brooks-Gunn and Duncan 1997; Cunha and Heckman 2007). However, there may be significant indirect effects of parental attention that operate by increasing the likelihood that youth will engage in delinquency or otherwise endanger their educational future in the high school years. Older children in large sibships may be especially at risk because parents need to pay closer attention to younger siblings once older children are moderately self-sufficient. Additionally, older children (especially females) may be called upon to participate in child care for younger siblings, taking away time that might otherwise be spent studying or engaged in education-related activities.

The potential importance of family size for educational and occupational attainment has been examined in a number of academic contexts. It was often included in early status attainment models, retaining a negative and significant relationship to occupational status once other demographic factors had been taken into account (Blau and Duncan 1967; Featherman and Hauser 1978). However, it was generally treated as a control variable in these works, and not being of particular interest in and of itself (Downey 1995). Although theories of why the number of siblings in a family had a causal effect on intelligence and/or educational attainment were proposed relatively early, the preponderance of evidence for the relationship did not accrue until the 1980s with the work of Blake, along with that of Steelman and colleagues (Blake 1981; Blake 1985; Mercy and Steelman 1982; Steelman 1985; Steelman and Mercy 1983). Blake's 1989 book served as the capstone to this line of research, analyzing multiple national datasets and documenting a consistent and strong negative relationship between sibship size and achievement in school (Blake 1989). While researchers believed that the broad empirical evidence for a negative effect of sibship size was

strong, the direct evidence for the mechanisms through which the effect operated was not convincingly demonstrated.

As noted above, the alternative to the claim that there is a causal relationship between sibship size and educational attainment, whether the mechanism is via intellectual environment or resource dilution, is that the relationship is caused by some other characteristic of parents or families that is associated with having more children and with children receiving less education. This paper will critically examine three possible factors that could produce the appearance of a sibship size penalty when there is in fact no causal relationship between family size and educational attainment.

The first potential factor that could produce a spurious relationship is parental education. It has been extensively documented that parents with more education have children who obtain more education (Blau and Duncan 1967). Because parents with more education also have fewer children on average, this relationship alone could explain the relationship between sibship size and educational attainment (US-Census 2012). However, past model-based examinations of the sibship size penalty have adjusted for parental education and found that it reduces, but does not eliminate, the sibship size penalty (Blake 1985). Therefore, past work does not seem to indicate that the entire sibship size penalty is due to parental education. Nevertheless, the analysis in this paper will adjust for parental education to remove the portion of the sibship size penalty that is due to this factor.

Another potential factor that could underlie the relationship between sibship size and educational attainment is parental religion. Because the average number of children per family varies by religion, if children in families who follow religions with larger families also receive less education this could induce a spurious relationship between sibship size and educational attainment (Blake 1989). It is also possible that the intensity of religious belief, or religiosity, could be associated with family size and

child educational attainment (regardless of the particular religion in which a parent believes). To our knowledge, religion and religiosity have not been examined as potential confounders for the relationship between sibship size and educational attainment. This paper's analysis will evaluate that possibility.

A third factor that could induce a spurious sibship size penalty is family structure. One type of family structure could clearly increase the number of siblings drawing from a resource stream: remarriages of divorced parents with children lead to step-families that have more children than either of the original families. (However, resources from parents outside the step-family could alleviate the degree to which economic resources are diluted.)

These explanations lead to a number of predictions for the empirical analysis to follow. First, the most basic prediction: sibship size will be negatively related to 4-year college enrollment and completion. Second, the sibship size penalty will remain even when the three potential factors that could generate a spurious relationship between family size and college enrollment and completion - parental education, parental religion, and family structure - are adjusted for. Third, because the effect of sibship size may operate via academic/intellectual ability, the sibship size effect will decrease once these factors have been adjusted for in the model. To our knowledge, researchers have not empirically examined whether the relationship between sibship size and educational attainment varies by race/ethnicity. Therefore, we make no predictions about racial differences in the sibship size penalty.

Data and Methods

This examination of the role of sibship size in racial disparities in educational attainment analyzes data from a large, nationally-representative survey conducted by the U.S. Bureau of Labor Statistics: the National Longitudinal Survey of Youth 1997

(NLSY97). This survey contains extensive and detailed data about each respondent's family, living situation, and educational experiences.

The NLSY97 was a probability sample of youth ages 12-16 in 1997, with additional surveys annually since that time. This paper analyzes data from 1997 to 2009. The full sample is comprised of two sub-samples, cross-sectional and ethnic, both of which will be used for the analysis. Only respondents remaining in the sample after 13 waves are included in the analysis, yielding a final analytic sample size of N=5,364.

The primary variable of interest for the analysis is "sibship size" – the number of siblings drawing from the family's resource pool, including the respondent. To reflect the complexity of family structure, the sibship size variable counts all biological, step, adopted, and foster siblings living in the household in the first wave of the survey. It also counts full biological siblings, half biological siblings who share the same father, and adopted siblings living outside the household, since they were presumably household members during a respondent's developmental period and/or drew from the same resource pool. 3.2% of sample respondents (an estimated 2.5% of youth aged 12-16) had more than 5 siblings present in the household, and youth from these very large families had a low probability of obtaining higher education based on observable characteristics other than family size such as income and parents' education. Because of this, the sibship size variable was top-coded at 5 to ensure that any observed sibship size effects are not being disproportionately driven by these largest family youths' particularly low *a priori* probability of obtaining higher education. Because this top-coding restricts the variance of the sibship size, the model findings for sibship size are necessarily conservative – the regression slopes will be uniformly lower than they would be without this adjustment.

This analysis also includes two other aspects of sibling composition that past research indicates may be relevant for the relationship between sibship size and educational attainment (de Haan 2010; Powell and Steelman 1990). The first aspect is whether or not siblings are closely spaced. The resource dilution explanation for the sibship size penalty posits that for a family's fixed amount of income and parental availability, additional children dilute the amount available for each child. If costs and need for attention vary with time, closely spaced children will require more resources at the same time than would less closely spaced children. In the context of higher education, the primary impact of closely spaced children would occur when more than one child is in college at a time, thus requiring more resources than if only one child was in college. Therefore, this analysis operationalizes "closely spaced" as having another sibling who was born within 3 years of a given respondent because that is the maximum spacing in which children could simultaneously attend college. The second aspect of sibling composition that may be relevant to the relationship between sibship size and educational attainment is birth order. There are a number of reasons birth order could impact educational attainment, and many have been proposed in the literature (Steelman et al. 2002). In the context of higher education, the most likely mechanism through which birth order would affect educational attainment may be via resource depletion over time. For example, a family's savings to pay for college may be depleted by older siblings, leaving younger siblings with greater financial burden and potentially forcing them to absorb more college cost via loans. Operationalizing birth order is difficult, however, because it is highly collinear with sibship size (e.g., one cannot have a birth order of "third" without having a sibship size of 3). Therefore, this analysis adopts the approach of Booth and Kee, who create a birth order index that normalizes birth order by the number of siblings present in the household (Booth and Kee 2008). Specifically, the index value for a given respondent is calculated as his

birth order divided by the average birth order in the family. This means that higher values for the index are associated with younger age because someone born third is younger than someone born second. Because this analysis top-codes sibship size at 5, the index ranges from 1/3 to 1-2/3 in the models below.

The primary outcome variables of interest for this analysis are 4-year college enrollment and 4-year college completion. This paper does not analyze 2-year college attendance due to (1) the historical U.S. policy focus on 4-year college attendance over 2-year college attendance, (2) the large increase in 2-year college attendance over the time periods, (3) the relatively low cost of 2-year colleges relative to 4-year colleges (which implies that the likelihood of cost constraints on attendance are lower), and (4) the additional analytical complexity added by analyzing which youth attend which type of college. Of course, attending any form of higher education is encouraged by the U.S. government whose programs serve both college types. However, the greatest returns to education are concentrated among 4-year college degrees which are currently of primary importance with respect to educational inequality in the U.S. population.

One particularly valuable piece of information that was collected in the NLSY97 is respondents' scores on the Armed Services Vocational Aptitude Battery (ASVAB) test. Direct evidence of intellectual ability is rare in surveys of this size and representativeness. The arithmetic reasoning and paragraph comprehension scale scores from this test are included in some of the estimated models below to adjust for differences in academic ability. Following NLSY documentation recommendations, the scaled versions of the scores are used which allows comparisons across individuals (National Longitudinal Surveys 2013a).

As noted above, family structure may be an alternative explanation for racial differences in the relationship between sibship size and educational attainment.

Generally, family structure consists of who is present in a child's household and their relationships with one another – a married father and mother along with their children in a "traditional" family; a single mother (either divorced or never-married) in some families; and a biological parent and step-parent in other families. By definition, children have a family structure at all times from birth to age 17 (after which many leave home to live at college, with roommates, or on their own). Over these 17 years, many children will live in a household where family structure changes, perhaps due to divorce or remarriage. Therefore, a child whose family experiences a divorce has two family structures: he lived with both biological parents before the divorce and in some other configuration or configurations after. How should this temporal variation be taken into account in measurement of family structure? If one measures family structure as the family configuration observed during a single measurement, one will over-simplify the child's family history (and may observe a family configuration that the child only experienced briefly). One could also define family structure in terms of whether a child's parents had undergone a marital transition or not. This definition would capture more information, but would not distinguish a child whose biological parents had divorced and remained unmarried from one whose parents were re-married (and perhaps brought other children into the marriage). One could also measure family structure as the configuration in which a child spent the most time; however, this definition would fail to capture the number of family structure transitions a child had experienced.

Family scholars have recognized this complexity in their examinations of family structure and have concluded that both the number and type of family transitions are important components of the effect of family structure on a child's life (Coleman, Ganong and Fine 2000; Fomby and Cherlin 2007; Wu and Martinson 1993). Each family transition requires a child to negotiate his or her relationships with

the other, perhaps new, family members in the household. Some research has suggested that many children adjust in the wake of these transitions in two or three years (Hetherington 1992). This implies that most family structure transitions do not permanently affect children. However, it also implies that serial family transitions leave children in a state of continuous familial flux which may impact their performance in school or their development more generally (depending on age). This paper operationalizes family structure by combining the two metrics of family history and the number of family transitions a youth experienced.

Of note, cohabitation has become an important component of family formation and composition, increasing significantly between the late 20th century and early 21st century (Bumpass and Lu 2000). In 2012, 3.6% of children lived with one biological parent and their cohabiting partner (US-Government 2013a). However, rates of cohabitation were lower in the 1980 to 2002 window during which NLSY97 respondents were 17 years old or younger. In the NLSY97 data, cohabiting partnerships were only observed after the survey began, whereas marriages were observed from the child's birth onwards. Including the observed cohabiting partnerships in the measurement of family structure history might bias inferences about cohabitation because only respondents whose parents cohabited while they were ages 12 to 17 would be identified, and they may not represent a random subset of respondents whose biological parents ever cohabited. Due to these measurement issues, only marriage relationships are included in the family structure history measure.

All analyses use panel sample weights to make the data representative of the population from which the survey was drawn. Only respondents remaining in the sample after 13 waves were included in the estimates, but the weights are designed to maintain the sample's representativeness.

All item-missing data for model variables were singly imputed. For continuous variables, a regression model on the outcome was estimated among respondents with data, and the values for those missing data were imputed based on predicted values from the regression model. For binary variables, a logistic regression model with the variable as the outcome was estimated among those with missing data, and the values for respondents missing the variable were imputed based on the predicted probabilities from the model. For categorical variables, a hot-deck procedure was used to randomly select a variable value from sample members who matched the individual with missing data on a number of characteristics. The characteristics were race/ethnicity, whether their residence was in a rural area, whether they attended public or private school, whether they graduated from high school, whether they enrolled in a 4-year college, and whether they graduated from a 4-year college.

After presenting descriptive statistics and bivariate trends, the general strategy for the analysis is to estimate regression models and examine the marginal slopes of the sibship size variable by race subpopulations. The focus of the analysis will be the patterning of and marginal relationship between sibship size and educational attainment across the subpopulations.

Descriptive Statistics

Tables 1 and 2 display summary statistics for the survey. About three quarters of the population is White or Other race, with both Black and Hispanic youth comprising approximately 14% each. About 10% of youth are only children. 35% of youth have one other sibling, about 30% have two other siblings, and 10 to 15% have 3 or 4 siblings. 5% of the youth population are first generation immigrants, and 10% are second generation immigrants. About 30% of respondents live in rural areas. The vast majority of youth age 12-16 attend public schools, and a slight majority attend

Table 1 - Summary Statistics - National Longitudinal Survey of Youth 1997

Variable	Values	Proportion	SE
Race/Ethnicity	White/Other	0.72	0.007
	Black	0.15	0.005
	Hispanic	0.13	0.005
Gender	Male	0.51	0.008
	Female	0.49	0.008
Number siblings of respondent	0	0.10	0.005
	1	0.35	0.008
	2	0.29	0.008
	3	0.15	0.006
	4	0.10	0.005
Immigrant Generation	1	0.05	0.003
	2	0.10	0.005
	3+	0.85	0.006
Region	Northeast	0.18	0.007
	North Central	0.28	0.008
	South	0.34	0.008
	West	0.20	0.007
Rural	Yes	0.28	0.008
	No	0.72	0.008
School type	Private/Parochial	0.10	0.005
	Public	0.90	0.005
Student-teacher ratio	<14	0.23	0.007
	14 to <18	0.40	0.008
	18 to <22	0.22	0.007
	22+	0.14	0.006
School size	<100	0.01	0.001
	100-299	0.07	0.004
	300-499	0.13	0.006
	500-749	0.24	0.007
	750 to 999	0.18	0.006
	1000+	0.38	0.008
Retained in grade	Yes	0.14	0.005
	No	0.86	0.005
Parent Religion	Catholic	0.32	0.008
	Mainstream Protestant	0.61	0.008
	Other Religion	0.04	0.004
	No Religion	0.03	0.003
Combined Family History	Both married bio parents since birth	0.48	0.009
	Bio parent since birth, 1 marital change	0.18	0.006
	Bio parent since birth, 2 marital changes	0.09	0.005
	Bio parent since birth, 3+ marital changes	0.09	0.005
	Single parent since birth	0.06	0.003
	Non-bio parent at some point	0.11	0.005

schools with student-teacher ratios below 18. A significant majority of youth attend schools with 500 or more students, and about 15% of youth have been retained in grade at some point. About 30% of youths' parents are Catholic, and about 60% are mainstream Protestant. 4% of parents follow another religion, and 3% of parents follow no religion. About half of youth live in "traditional" families with both biological parents since birth. About 20% of youth live with a biological parent who has experienced one marital transition; approximately 10% of youth have experienced 2 family transitions, and another 10% have experienced 3 or more family transitions. 6% of youth have lived with a single, never married parent since birth, and 11% of youth have lived without a biological parent present at some point.

Table 2 provides additional descriptive statistics about the population of youth aged 12-16 in 1997. Non-GED high school completion is approximately 80%, consistent with the findings of Heckman and LaFontaine (2010). Approximately 20% of respondents were born to unmarried parents. Mothers were an average of 26 years old when respondents were born, and they were an average of 23 years old when they had their first child. Mothers obtained an average of slightly more than 13 years of

Table 2 - Summary Statistics - National Longitudinal Survey of Youth 1997

Variable	Mean	SE
Graduate HS	0.83	0.006
ASVAB Paragraph Comprehension Scale Score	-166.16	14.120
ASVAB Math Scale Score	-257.88	14.855
Age	14.68	0.021
Nonmarital Birth	0.21	0.006
Mother's age at R's birth	25.86	0.086
Mother's age at first birth	23.35	0.078
Mother's Highest Grade	13.13	0.046
Father's Highest Grade	12.92	0.051
Annual income (thousands of 1990 dollars)	42.19	0.596
Parent religiosity	370.94	2.502

education, while fathers received an average of slightly less than 13 years. The average family annual income was about $42,000 1990 dollars. The parent religiosity index, which measures the frequency of religious behaviors, has a range of 0 to 600 with a mean of 371.

Table 3 displays mean sibship sizes at the household level by demographic characteristics.[2] The overall mean sibship size in the population is about 2.7 children per family, consistent with research on changing U.S. family sizes over the 20th century (Wetzel 1990). White and other race population members have families with an average of about 2.6 children, black families have larger families with about 2.8 children, and Hispanic families have the largest families with approximately 3 children. As might be expected based on Catholic beliefs about birth control, Catholic

Table 3 - Mean Sibship Size by Demographics – National Longitudinal Survey of Youth 1997

Variable	Values	Mean	SE
Overall		2.66	0.019
Race	White/Other	2.58	0.023
	Black	2.77	0.043
	Hispanic	2.97	0.049
Parent Religion	Catholic	2.75	0.033
	Mainstream Protestant	2.62	0.024
	Other Religion	2.62	0.107
	No Religion	2.55	0.118
Combined Family History	Both married bio parents since birth	2.68	0.026
	Bio parent since birth, 1 marital change	2.65	0.045
	Bio parent since birth, 2 marital changes	2.66	0.064
	Bio parent since birth, 3+ marital changes	2.63	0.072
	Single parent since birth	2.60	0.079
	Non-bio parent at some point	2.61	0.060
Non-marital birth	Yes	2.67	0.043
	No	2.65	0.021
Income Quartile	1 Lowest	2.87	0.046
	2	2.76	0.041
	3	2.59	0.037
	4 Highest	2.52	0.031

[2] This and all subsequent analyses of sibship size are calculated using sibship size top-coded at 5, as described in the previous section.

parents have an average of about .13 more children than do parents of other religions or no religion. Average family size does not appear to vary by family structure history, with families of all types having about 2.6 children. This is somewhat unexpected, as we anticipated that families who had experienced divorce and remarriage would have larger families due to step-children. However, there may have been few enough of these families in 1997 for the NLSY97 not to have the statistical power to detect the relationship. Youth born to non-married parents do not have more siblings than youth born to married parents. Table 3 also demonstrates that mean sibship sizes decrease as income increases; the difference in sibship sizes between the bottom and top income quartiles is approximately .35 children.

Table 4 presents parent religion, combined family history, non-marital birth rates, and income quartile by race/ethnicity. About 30% of white and other race parents were Catholic; in contrast 7% and 77% of black parents and Hispanic parents were Catholic, respectively. About 60% of white and other race parents were mainstream Protestant, as compared to an overwhelming majority of 91% of black parents and about 20% of Hispanic parents. Rates of following another religion or no religion were highest among white and other race parents, at about 4%; very small proportions of black and Hispanic parents were not Catholic or mainstream Protestant.

Among families with white or other race youth, 55% lived with both biological parents since birth, a greater proportion than the 48% overall. 17% of youth experienced one marital change during their childhood, and 9% experienced two and three. Only 3% of white and other race youth lived with a single parent for their entire childhood, and 8% lived with non-biological parents at some point. In contrast to white and other race youth, 19% of black youth lived with both married biological parents from birth to age 17. Similar proportions of black youth experienced one, two, and three marital changes as white and other race youth. Significantly more black

youth lived with a single parent throughout their childhood – 18%, compared to 3% of white and other race youth (although these single parents could have had cohabiting partners). A significantly greater proportion of black youth than white and other race youth lived without their biological parents at some point – 24%, compared to 8%. Finally, Hispanic youth have similar family histories to white and other race youth, with slightly lower rates of 2 and 3+ marital changes and moderately higher rates of living without biological parents at some point.

Table 4 - Family Characteristics by Race/Ethnicity - National Longitudinal Survey of Youth 1997

		White/Other		Black		Hispanic	
Variable	**Value**	**Prop.**	**SE**	**Prop.**	**SE**	**Prop.**	**SE**
Parent Religion	Catholic	0.29	0.010	0.07	0.009	0.77	0.017
	Mainstream Protestant	0.62	0.011	0.91	0.010	0.21	0.017
	Other Religion	0.05	0.005	0.01	0.003	0.00	0.003
	No Religion	0.03	0.004	0.01	0.004	0.02	0.005
Combined Family History	Both married bio parents since birth	0.55	0.011	0.19	0.013	0.48	0.019
	Bio parent since birth, 1 marital change	0.17	0.008	0.21	0.013	0.19	0.014
	Bio parent since birth, 2 marital changes	0.09	0.006	0.10	0.009	0.07	0.009
	Bio parent since birth, 3+ marital changes	0.09	0.006	0.08	0.009	0.05	0.007
	Single parent since birth	0.03	0.003	0.18	0.012	0.09	0.010
	Non-bio parent at some point	0.08	0.006	0.24	0.013	0.12	0.012
Non-marital birth		0.13	0.007	0.56	0.017	0.22	0.016
Income Quartile	1 Lowest	0.11	0.007	0.39	0.016	0.33	0.018
	2	0.20	0.009	0.31	0.016	0.31	0.018
	3	0.30	0.010	0.19	0.014	0.20	0.016
	4 Highest	0.40	0.011	0.10	0.010	0.16	0.014

There are significant differences by race in the proportion of births to unmarried parents. 13% of white and other race youth were born to unmarried parents, as compared to 56% among black youth; Hispanic youth fall in the middle with about 22% non-marital births. Income distributions vary greatly by race as well, with 70% of

white and other race youth's families falling in the top half of the income distribution. In contrast, only 30% of black youth's families have earnings in the top half of the income distribution. Hispanic families have slightly higher average incomes than black families, with 36% falling in the top half of the income distribution.

Table 5 displays the college outcome variables of interest by number of siblings, race, parent religion, combined family history, non-marital birth, and income quartile. Table 5 replicates a pattern described in past literature on the relationship between sibship size and educational attainment – instead of a monotonic negative relationship between the two, we observe that youth from single child families are less likely to enroll or complete than are youth from two-child families (Steelman et al. 2002). Among youth with one or more siblings, however, there is a net negative relationship between sibship size and 4-year enrollment, where 57% of children from the two child families enroll, as compared to 32% of children from the largest families. The same non-monotonicity is observed for 4-year college completion. 31% of youth with no siblings as compared to 39% of youth with one sibling graduate from college. As with enrollment, rates of college completion decline as families have more siblings. This bivariate relationship between sibship size and educational attainment reflects marginal differences; the models below examine whether this gradient remains once other potential explanations than sibship size have been taken into account.

Reflecting the findings of the research described above, there are significant differences among racial/ethnic groups in higher education attainment. Just over half of white and other race youth enroll in 4-year college, while only 35% of black and Hispanic youth enroll. Differences in graduation from 4-year college are larger in magnitude; white and other race youth are more than twice as likely as black and

Table 5 – 4-Year College Enrollment and Completion by Demographics – National Longitudinal Survey of Youth 1997

		4-year Enrollment		4-year Completion	
Variable	**Values**	**Proportion**	**SE**	**Proportion**	**SE**
Overall		0.48	0.008	0.32	0.008
Number siblings	0	0.48	0.024	0.31	0.022
	1	0.57	0.014	0.39	0.014
	2	0.46	0.015	0.33	0.015
	3	0.41	0.022	0.25	0.019
	4+	0.32	0.023	0.19	0.021
Race	White/Other	0.53	0.010	0.38	0.010
	Black	0.36	0.015	0.17	0.012
	Hispanic	0.34	0.017	0.18	0.014
Parent Religion	Catholic	0.49	0.014	0.33	0.013
	Mainstream Protestant	0.46	0.010	0.31	0.010
	Other Religion	0.70	0.039	0.52	0.044
	No Religion	0.44	0.051	0.29	0.046
Combined Family History	Both married bio parents since birth	0.62	0.011	0.46	0.012
	Bio parent since birth, 1 marital change	0.43	0.018	0.25	0.016
	Bio parent since birth, 2 marital changes	0.37	0.025	0.22	0.022
	Bio parent since birth, 3+ marital changes	0.32	0.025	0.17	0.020
	Single parent since birth	0.28	0.026	0.14	0.021
	Non-bio parent at some point	0.27	0.020	0.13	0.015
Non-marital birth	Yes	0.29	0.014	0.14	0.011
	No	0.53	0.009	0.37	0.009
Income Quartile	1 Lowest	0.22	0.013	0.09	0.010
	2	0.32	0.015	0.18	0.013
	3	0.49	0.016	0.30	0.015
	4 Highest	0.74	0.013	0.58	0.015

Hispanic youth to complete, with 38% and 18% completion rates, respectively. Educational attainment rates also vary by parent religion. All groups enroll at about the overall average rate of 48% except for those of "Other" religion, who enroll at the higher rate of 70%. Similarly, the other religion youth are significantly more likely to graduate from college, with a 52% rate compared to about 30% among Catholics, mainstream Protestants, and those with no religious identification. The other religion group is primarily comprised of youth with parents who identify as Jewish.

Table 5 also presents rates of educational attainment by combined family history, including high school graduation, enrollment in 4-year college, and graduation

from 4-year college. As one might anticipate from previous research, youth who had "traditional" married biological parent families attain education at rates above the overall average. In contrast, youth from all other family histories attain education at rates below the overall average. Enrollment in 4-year college monotonically decreases as one moves from youth experiencing the stability of traditional families to the relative instability of families that experienced three or more marital transitions, with enrollment rates of 62% and 32%, respectively. 28% of youth from single parent families enroll in four-year college, and a similar proportion of youth enroll who lived with non-biological parents at some point. A similar pattern is present for graduation from 4-year college, with the highest proportion of graduating youth from traditional families and the lowest proportion of graduating youth living without a biological parent present at some point. There are also differences in educational attainment by whether a respondent was born to unmarried parents. Youth born to married parents are about 65% more likely to enroll in 4-year college than are youth born to unmarried parents. The disparity in 4-year college completion is even greater, with youth born to married parents over 250% more likely to complete college than youth born to unmarried parents. This is evidence of significant attrition after college enrollment among youth born to unmarried parents.

As recently documented, there is a steep and increasing gradient in college enrollment and completion by income (Bailey and Dynarski 2011; Reardon 2011). For enrollment, there is a difference of 52 percentage points between the highest and lowest income quartiles. Similarly, for completion there is a difference of 49 percentage points between the highest and lowest income quartiles.

Overall, Table 5 paints a detailed demographic picture of 4-year college attendance and completion of youth ages 12-16 in 1997. Higher educational attainment is unevenly distributed among the population, with the highest rates among

those with higher incomes, those who are white or other race, those born to married parents, and those who have lived with both biological parents since birth. In the next section, we turn to analysis of the relationship between sibship size and educational attainment and the competing explanations for why such a relationship exists.

Regression Models

This section of the paper estimates regression models of the relationship between sibship size and higher educational attainment, focusing on alternative explanations for the negative marginal relationship described above. As observed in past research and the descriptive statistics already presented, the relationship between sibship size and educational attainment is non-monotonic (Steelman et al. 2002). It is unclear why there is an increase in educational attainment between youth with no siblings and youth with one sibling, and past research has not examined that relationship in depth; we leave this topic to future research. To reflect the non-monotonicity of the relationship between sibship size and education, the regression models below fit a piecewise linear parameterization of sibship size. Specifically, sibship size is decomposed into two parts: one term for the difference between youth with no siblings and those with one sibling, and a linear term for sibship sizes from those with one sibling to those with four or more siblings. The latter term will be the focus of the analysis. The tables below display a series of regression models with the same outcome variable, adding explanatory variables to the models that reflect the competing explanations for the sibship size penalty. Table 6 presents the estimated models for enrollment in 4-year college.

The first model in Table 6 includes only variables reflecting sibship size and composition. Reflecting the descriptive statistics above, it indicates that there is an increase in educational attainment from youth with no siblings to youth with one

sibling when sibship spacing and birth order have been included in the model. Additionally, there is a significant negative relationship between having more than one other sibling and college enrollment, with the odds of enrolling decreasing by 30% for each additional sibling beyond one. Neither close spacing of siblings nor birth order are related to enrolling in 4-year college. The second model in Table 6 adds race/ethnicity and interactions between race/ethnicity and sibship size. Reflecting racial differences in college enrollment, black and Hispanic youth are less likely to enroll than are white and other race youth. As in the first model, having one sibling is associated with higher rates of enrollment than having no siblings. The negative relationship between having more than one sibling and college enrollment remains significant and of similar magnitude, and any differences in the relationship between sibship size and enrollment by race are too small to be detected with this sample.

The third model adds variables corresponding to one potential explanation for the sibship size penalty: religious preference and religiosity. As described above, religion could explain the relationship between sibship size and education if persons of certain religions (or who are highly religious) are likely both to have more children and to have children with lower educational attainment. The third model indicates that youth whose parents are Catholic have odds of enrolling that are about 30% greater than those for mainstream Protestants. Youth whose parents follow a religion other than Catholicism and mainstream Protestantism have odds of enrolling in college almost three times greater than those for youth whose parents are mainstream Protestants. Additionally, parent religiosity is positively related to enrollment in 4-year college. It is important to note that this relationship is not the result of the racial composition of different religions because race/ethnicity is included in the model. The race coefficients have similar magnitudes to those in model 2. Model 3 indicates that parent religion and religiosity do not explain the relationship between sibship size and

Table 6 - 4-Year College Enrollment Regression Models - National Longitudinal Survey of Youth 1997

Variable	Variable Value	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Sibship size 1 to 2		1.316*	1.272*	1.271*	0.994	1.062	1.17
		0.149	0.146	0.148	0.122	0.144	0.176
Sibship size >2		0.707***	0.759***	0.732***	0.774***	0.847**	0.857*
		0.025	0.036	0.035	0.038	0.049	0.052
Sibling within +/- 3 years		1.096	1.068	1.064	1.029	1.031	0.949
		0.081	0.08	0.08	0.079	0.088	0.089
Sibling order index		1.051	1.057	1.043	0.841	0.737	0.951
		0.088	0.089	0.088	0.075	0.099	0.136
Race/Ethnicity	Black		0.578***	0.580***	1.09	1.313*	3.318***
			0.061	0.064	0.133	0.177	0.541
	Hispanic		0.554***	0.497***	0.573***	1.057	1.397
			0.069	0.065	0.08	0.174	0.267
Sibship Size X Race/Ethnicity							
	Black		0.942	0.98	0.908	0.987	0.999
			0.072	0.075	0.072	0.083	0.09
	Hispanic		0.897	0.912	0.853	1.023	1.005
			0.077	0.078	0.078	0.102	0.109
Parent Religion	Catholic			1.285**	1.216*	1.108	1.146
				0.105	0.101	0.101	0.114
	Other Religion			2.815***	2.534***	1.511	1.732*
				0.567	0.484	0.33	0.398
	No Religion			0.884	0.856	0.821	0.693
				0.193	0.2	0.214	0.183
Parent Religiosity Scale				1.001*	1.000	1.000	1.000
				<.001	<.001	<.001	<.001
Non-marital Birth					0.647***	0.911	1.024
					0.072	0.111	0.138
Combined Family History	Bio parent since birth, 1 marital change				0.504***	0.718**	0.759*
					0.048	0.078	0.091
	Bio parent since birth, 2 marital changes				0.383***	0.464***	0.490***
					0.049	0.063	0.07
	Bio parent since birth, 3+ marital changes				0.286***	0.416***	0.437***
					0.037	0.061	0.068
	Single parent since birth				0.393***	0.605**	0.546**
					0.068	0.114	0.115
	Non-bio parent at some point				0.276***	0.406***	0.435***
					0.036	0.056	0.07
+Demographics & School Vars						√	√
+ASVAB & Retained in Grade							√
N		5364	5364	5364	5364	5364	5364
Pseudo-R^2		0.018	0.033	0.041	0.093	0.219	0.335

legend: * p<0.05; ** p<0.01; *** p<0.001

enrollment in 4-year college. The positive relationship between having one sibling rather than no siblings and enrollment remains, and the negative relationship between having more than one sibling and enrollment also remains at similar magnitudes to the earlier models.

Model 4 adds variables representing another potential explanation for the relationship between sibship size and educational attainment: family structure and non-marital birth. These factors could explain the sibship size penalty if parents who have children while unmarried, parents who divorce, or parents who never marry have larger families and have children who are less likely to enroll in 4-year college. Reflecting the descriptive statistics described above, the odds of enrollment among youth born to unmarried parents are about 35% less than those of youth born to married parents. Youth from all family structures other than both biological parents have significantly lower odds of enrolling than do youth who have lived with both biological parents since birth. The magnitude of the differences by family structure is quite large, ranging from a decrease in odds of enrolling of 50% for youth who experienced one family transition to decreases of over 70% for youth who experienced three or more family transitions or lived with a non-biological parent at some point. The relationships between religion and enrollment are of similar magnitudes as in model 3; religiosity is no longer significant. Hispanic youth are significantly less likely to enroll than white and other race youth, while black youth are not less likely to enroll. There is not support for the claim that family structure or non-marital births underlie the negative relationship between sibship size and enrollment in 4-year college. After adjusting for family structure and non-marital birth, the coefficient for number of siblings greater than one is still significant and negative, with a 23% decrease in the odds of enrollment for each additional sibling a youth has beyond one.

Model 5 adds variables representing demographics and school characteristics to the model. These variables include family income and the levels of education attained by respondents' parents which are both known to be strongly related to enrollment in college. The inclusion of these variables reveals that many of the relationships observed in model 4 were at least partially due to demographic composition. The relationships between religion and enrollment observed in models 3 and 4 have now lost significance, although youth whose families follow a religion other than Catholicism or mainstream Protestantism appear to be more likely to enroll. Being born to unmarried parents is no longer associated with enrollment in college, but all family structure histories remain associated with lower enrollment rates than those of youth who lived with both biological parents since birth. Once demographic factors have been taken into account, Hispanic youth are not less likely to enroll than are white and other race youth. The coefficient for black youth is now significant and positive, indicating that the odds of enrolling in 4-year college are about 30% higher than a white youth attending comparable schools and having similar demographic and income characteristics. This finding replicates previous research findings that black youth are more likely to attend college than white youth in models that include an adjustment for income (Black and Sufi 2002). The interactions between sibship size and race remain non-significant. Even when income and parental education have been included in the model the negative relationship between sibship size and the probability of enrollment in 4-year college remains significant. However, its magnitude has decreased with each additional sibling beyond one associated with a decrease of 15% in the odds of enrollment. The proportion of the variance in college enrollment that model 5 explains has more than doubled relative to model 4, with a pseudo-R^2 of .219. The sixth model adds the ASVAB test score variables and whether the respondent was ever retained in grade; this model is a direct test of the confluence

explanation because decreased intellectual ability would presumably be reflected in test scores and school performance. It largely shows the same coefficient patterns observed in model 5, including a negative relationship between sibship size and enrollment of similar magnitude. This implies that academic ability does not play a strong mediating role for the effect of sibship size on 4-year college enrollment and, therefore, that the confluence explanation does not provide an accurate description of the mechanism through which the sibship size penalty operates. While model 6 shows a strong negative relationship between not living with both biological parents since birth and enrollment in college, it does not support the contention that family structure is the true underlying cause of the sibship size penalty. Additionally, model 6 does not indicate that there is a relationship between being born to unmarried parents and enrollment once other factors have been adjusted for, so it does not support the claim that non-marital fertility is the true underlying cause of the sibship size penalty. The proportion of the variance in 4-year college enrollment that model 6 explains increased by over 50% relative to model 5; it explains about one third of the variance, which is reasonably large for models of educational outcomes.

Table 7 displays the regression models examining completion of 4-year college. The first model includes variables representing sibship composition, the piecewise linear function for sibship size, whether or not siblings were closely spaced, and the birth order index. Mirroring the patterns described above, the coefficient for having one sibling versus having no siblings is positive and significant; the odds of youth with one sibling completing 4-year college are about 28% greater than the odds for youth with no siblings. As with enrollment, having more than one sibling is negatively related to college completion, with the odds of completion decreasing by about 30% for each additional sibling beyond one. Surprisingly, having a closely spaced sibling is positively associated with completing 4-year college. It is possible

Table 7 - 4-Year College Completion Regression Models - National Longitudinal Survey of Youth 1997

Variable	Variable Value	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Sibship size 1 to 2		1.278*	1.224	1.229	0.938	1.02	1.117
		0.157	0.154	0.156	0.127	0.153	0.182
Sibship size >2		0.703***	0.776***	0.756***	0.804***	0.891	0.896
		0.029	0.04	0.038	0.042	0.055	0.056
Sibling within +/- 3 years		1.224*	1.183*	1.183*	1.149	1.162	1.101
		0.098	0.096	0.096	0.096	0.109	0.109
Sibling order index		1.109	1.119	1.109	0.904	0.87	1.16
		0.102	0.105	0.104	0.089	0.126	0.174
Race/Ethnicity	Black		0.453***	0.466***	0.904	1.077	2.420***
			0.054	0.058	0.125	0.166	0.437
	Hispanic		0.414***	0.380***	0.427***	0.776	0.976
			0.06	0.057	0.069	0.141	0.198
Sibship Size X Race/Ethnicity	Black		0.786*	0.809*	0.739**	0.805*	0.816
			0.075	0.078	0.074	0.085	0.095
	Hispanic		0.91	0.921	0.871	1.016	1.025
			0.094	0.095	0.096	0.125	0.136
Parent Religion	Catholic			1.240*	1.168	1.069	1.104
				0.108	0.104	0.104	0.114
	Other Religion			2.312***	2.055***	1.137	1.251
				0.436	0.379	0.226	0.285
	No Religion			0.858	0.826	0.78	0.666
				0.198	0.201	0.206	0.173
Parent Religiosity Scale				1.000	1.000	1.000	1.000
				<.001	<.001	<.001	<.001
Non-marital Birth					0.617***	0.928	1.074
					0.085	0.142	0.177
Combined Family History	Bio parent since birth, 1 marital change				0.465***	0.657***	0.692**
					0.049	0.08	0.091
	Bio parent since birth, 2 marital changes				0.348***	0.422***	0.465***
					0.05	0.064	0.074
	Bio parent since birth, 3+ marital changes				0.253***	0.378***	0.425***
					0.04	0.068	0.077
	Single parent since birth				0.404***	0.563*	0.493**
					0.09	0.131	0.124
	Non-bio parent at some point				0.235***	0.353***	0.370***
					0.038	0.059	0.067
+Demographics & School Vars						√	√
+ASVAB & Retained in Grade							√
N		5364	5364	5364	5364	5364	5364
Pseudo-R2		0.018	0.046	0.052	0.108	0.24	0.339

legend: * p<0.05; ** p<0.01; *** p<0.001

that this reflects family planning strategies in which parents intentionally have few closely spaced children. It is likely that parents who plan in this way would also make concerted efforts to ensure that their children were highly educated (Lareau 2011). The second model adds race/ethnicity and interactions between sibship size and race/ethnicity. The sibship composition variables have similar patterns to model 1. Both black and Hispanic youth have odds of completing 4-year college less than half of those for white and other race youth. The interaction between race and sibship size indicates that black youth experience a larger sibship size penalty than do youth of other races, with the odds of completing college decreasing about 20% more for each additional sibling beyond one a black youth has than for each additional sibling youth of other races have.

Model 3 adds variables representing parent religion and religiosity. Sibship composition, race/ethnicity, and the interaction between sibship size and race/ethnicity all have similar magnitudes as in model 2. Each additional sibling beyond one is associated with a 24% decrease in the odds of completing college, and black youth experience an additional 19% penalty. Parent religion is associated with the odds of completing; youth from Catholic families have odds of completing 24% greater than those of youth from mainstream Protestant families, and youth from families with religions other than Catholicism or mainstream Protestantism have odds of completing more than twice as large as those for youth from mainstream Protestant families. Because the negative relationship between sibship size and completion remains, the model does not indicate that parent religion underlies the sibship size penalty. Parent religiosity does not appear to be related to the likelihood of youth completing college.

Model 4 adds variables representing family structure and being born to unmarried parents, both of which could potentially explain the relationship between sibship size and college completion. Being born to unmarried parents is associated

with a 38% decrease in the odds of completing college. Having a family structure other than living with both biological parents since birth is also negatively related to college completion, with some family structures being associated with odds of completion 4 times smaller than those of youth from traditional families. Having parents that follow a religion other than Catholicism or mainstream Protestantism remains associated with higher odds of completing college, but having Catholic parents is no longer associated with completion. Hispanic youth still have odds of completion about half those of white and other race youth, but black youth are neither more nor less likely to complete. Sibship size remains negatively associated with completing 4-year college, and black youth experience a stronger penalty than do youth of other races.

Model 5 adds an assortment of demographic and school characteristic variables to model 4. The additions reveal that many of the associations observed in model 4 reflected demographic composition, with almost all variables other than family structure losing significance. Youth from some family structures are far less likely to complete college than others, with the odds of completion for youth who experienced three or more family transitions or lived with a non-biological parent at some point less than 40% of those for youth who lived with both biological parents since birth. While the sibship size penalty among white, Hispanic, and other race youth is no longer detectable, the penalty remains significant among black youth with a 20% decrease in the odds of completion for each additional sibling a black youth has beyond one. Model 6 adds variables representing academic ability – ASVAB test scores and whether the respondent was retained in grade. As observed in the enrollment models, black youth are significantly more likely to complete college than are white youth with similar demographic characteristics, who attend similar schools, who have similar family incomes, and who have similar academic ability. Family

structure remains significantly related to completion, but with the inclusion of academic ability none of the other variables of interest have detectably large relationships with completion. Sibship size no longer has a significant relationship to college completion for white or Hispanic youth, and it has a marginally significant relationship for black youth. This could indicate that much of the negative relationship between sibship size and college completion is mediated by academic ability.

Table 8 displays the marginal slope for sibship size across all 6 models from Tables 6 and 7. The average marginal relationship between sibship size and enrollment in 4-year college is -.084 on the probability scale, meaning that each additional sibling one has beyond one decreases the probability of enrolling in a 4-year college by .08. The strength of the relationship declines as adjustment variables are added to the model, and remains significant with a value of -.023 in the final model that adjusts for ASVAB scores and whether the respondent was retained in grade.

Table 8 - Sibship Marginal Slopes for 4-Year College Regression Models - National Longitudinal Survey of Youth 1997

	Enrollment			**Completion**		
Model	**Slope**	**SE**	**Sig.**	**Slope**	**SE**	**Sig.**
Sibship + Close/Order	-0.084	0.008	<.001	-0.075	0.008	<.001
+Race X Sibship	-0.071	0.009	<.001	-0.059	0.009	<.001
+Religion	-0.077	0.009	<.001	-0.063	0.009	<.001
+Family Structure	-0.063	0.009	<.001	-0.050	0.008	<.001
+Demographics/School Vars	-0.030	0.009	<.001	-0.022	0.008	0.009
+ASVAB & Retained in grade	-0.023	0.008	0.002	-0.017	0.007	0.017

The average marginal relationship between sibship size and completion is -.075, approximately the same magnitude as the relationship for enrollment. Again, the strength of the relationship declines as variables are added to the model, settling at a significant -.017 in the final model. However, these overall marginal relationships mask variation by race/ethnicity.

Table 9 displays the marginal sibship slope by race/ethnicity for the five models that include an interaction between sibship greater than 2 and race/ethnicity. Reflecting the odds ratios for sibship in the enrollment models, the sibship size marginal relationship is significantly negative for all race groups in the first model containing the interaction; each additional sibling a youth has beyond 1 is associated with a decrease of .07 in the probability of enrollment. The sibship size marginal relationship remains significant for all races and of similar magnitude in the model containing religion and religiosity. The addition of family structure and birth to unmarried parents decreases the magnitude of the marginal relationship for white youth, but it remains stable for black and Hispanic youth. Adding demographic and school variables makes the marginal sibship size relationship non-significant for Hispanics, implying that among Hispanic youth certain types of youth tend both to have more siblings than average and to be relatively unlikely to enroll in 4-year college; in other words, the sibship size relationship is unlikely to have a causal impact on the probability of enrollment among Hispanic youth. Once demographic factors have been included, the magnitude of the sibship size relationship decreases by about half for white, black, and other race youth. In the final model, ASVAB scores and whether the respondent was retained in grade are added. As in the previous model, the sibship size relationship is not significant for Hispanic youth. The relationship decreases in magnitude relative to the previous model for white, black, and other race youth; each additional sibling beyond one is associated with a decrease of .2 in the probability of enrollment in 4-year college. This means probability of enrollment for white, black, and other race youth from the largest families is about .07 smaller than the probability for youth with only one sibling. For white and other race youth, this corresponds to about a 14% decrease in the likelihood of enrollment relative to the

average probability of .53. For black youth, this corresponds to a 19% decrease in the probability of enrollment.

Table 9 - Sibship Marginal Slopes by Race/Ethnicity for 4-Year College Regression Models - National Longitudinal Survey of Youth 1997

		Enrollment			Completion		
Model	**Race**	**Slope**	**SE**	**Sig.**	**Slope**	**SE**	**Sig.**
Race X Sibship	White/Other	-0.068	0.011	<.001	-0.059	0.012	<.001
	Black	-0.076	0.014	<.001	-0.069	0.012	<.001
	Hispanic	-0.083	0.015	<.001	-0.050	0.013	<.001
+Religion	White/Other	-0.076	0.011	<.001	-0.064	0.011	<.001
	Black	-0.075	0.014	<.001	-0.068	0.012	<.001
	Hispanic	-0.087	0.015	<.001	-0.052	0.013	<.001
+Family Structure	White/Other	-0.057	0.011	<.001	-0.047	0.011	<.001
	Black	-0.074	0.013	<.001	-0.068	0.011	<.001
	Hispanic	-0.083	0.015	<.001	-0.048	0.013	<.001
+Demographics/School Vars	White/Other	-0.031	0.011	0.004	-0.020	0.011	0.060
	Black	-0.034	0.013	0.011	-0.039	0.011	<.001
	Hispanic	-0.024	0.014	0.100	-0.011	0.013	0.379
+ASVAB & Retained in grade	White/Other	-0.024	0.009	0.011	-0.016	0.009	0.077
	Black	-0.023	0.011	0.044	-0.030	0.010	0.003
	Hispanic	-0.021	0.013	0.119	-0.008	0.012	0.494

The pattern of the marginal relationships between sibship size and the probability of college completion by race are similar to those for enrollment. In the first model containing a sibship size interaction with race, the marginal relationships for youth of all races are about -.06. Adding religion and religiosity to the model has minimal impact on the estimated marginal relationships. Including family structure and birth to unmarried parents to the model leads to slight decreases in the sibship size marginal relationships, but they remain significant for youth of all races. As with enrollment, the additional demographic and school characteristics decrease the magnitude of the relationships by about half. Additionally, the relationships lose significance at the traditional 95% level for white, Hispanic, and other race youth. This implies that differences in the marginal probability of completion by sibship size for youth of those

races are largely due to demographic composition and probably not to a causal relationship. The sibship size relationship remains significant for black youth in the final model that includes ASVAB scores and whether the respondent was retained in grade, but the relationship is too small to be detected by this sample for white, Hispanic, and other race youth. This means that the relationship between sibship size and college completion may be causal for black youth, assuming no unobserved aspects of parents and families cause both larger families and decreased likelihoods of college completion. Black youth from the largest families have probabilities of college completion .09 smaller than black youth with only one sibling. This corresponds to a sizable decrease of 53% relative to the average probability of .17 among black youth and implies that sibship size may contribute to black-white disparities in rates of completing 4-year college.

Discussion and Conclusion

This paper has examined competing explanations for the sibship size penalty and whether the strength of the relationship between sibship size and 4-year college enrollment and completion differs by race.

The analysis did not find support for the claim that the relationship between sibship size and educational attainment is spurious. One potential factor that could generate a spurious relationship is parental religion; family size varies by religion, as does educational attainment. However, adjusting for parental religion and religiosity did not decrease the magnitude of the relationship between sibship size and college enrollment and completion. Another aspect of families that might create a spurious relationship between sibship size and educational attainment is family structure. Families that experience marital transitions or who have children outside of marriage may have more children and have children who are less likely to enroll in and

complete 4-year college. While the models revealed large and significant differences in enrollment and completion by family structure, its inclusion only slightly decreased the magnitude of the sibship size penalty. Thus, it does not appear that family structure underlies the sibship size penalty. A third factor that might create the appearance of a relationship between sibship size and educational attainment is parental education. Family size decreases as parental education increases, and parent and child levels of educational attainment are tightly linked. Therefore, it is possible that selection is operating; parents with higher education choose to have fewer children, making it appear as if there is a causal relationship between sibship size and educational attainment. The analysis finds moderate support for the selection hypothesis; the magnitude of the relationship between sibship size and enrollment and completion in higher education decreases by more than half when parental education and other demographic and school factors are adjusted for in the model. However, the sibship size penalty remains significant and of non-trivial magnitude - selection does not appear to tell the entire story. The confluence explanation for the sibship size penalty implies that the causal effects of sibship size operate by decreasing the intellectual ability of children from large families, and presumably intellectual ability would be at least partially reflected in academic achievement. Because the NLSY97 contains two measures of academic achievement, scores on standardized exams and whether the respondent was retained in grade, the analysis examined whether the sibship size penalty is likely to occur via this proposed causal mechanism. There was not strong support for this contention, with the relationship between sibship size and college enrollment and completion only slightly moderating once the models adjusted for academic achievement. However, the overall findings about the relationships among parental education, academic achievement, and sibship size mask important variation by race/ethnicity.

The relationship between sibship size and educational attainment was approximately equal across races/ethnicities for models adjusting for religion and family structure. However, the addition of parental education and other demographic characteristics revealed differences by race: the sibship size penalty for enrollment was present for both white and black youth, but it was not present for Hispanic youth. This implies that among Hispanic families parental selection likely generates the relationship between sibship size and educational attainment. However, the sibship size penalty for enrollment remained significant and of non-trivial size among white, black, and other race youth even when variables representing academic achievement were included. The variation by race/ethnicity differs for college completion. Once parental education and other demographic factors are included, the sibship size penalty only remains significant among black youth. The fact that a sibship size penalty is observed among white and other race youth for enrollment but not for completion implies that it does not operate via resource constraints due to having multiple children in college simultaneously; the absence of a negative relationship between closely spaced siblings and college completion also supports this contention. The relationship between sibship size and college completion among black youth does remain significant and non-trivial even with the inclusion of academic achievement. Therefore, it may play a small role in producing black-white disparities in college completion.

The analysis also replicated findings from earlier research that black youth are more likely to attend college than white youth once models have adjusted for demographics and income (Cameron and Heckman 2001; Hauser 1993). More recent work indicates that this difference was concentrated at the low end of the income distribution - among those with low income, black youth were more likely than whites to enroll, but among those with medium or high income, white youth were more likely

to enroll. It also finds that the higher probability of enrollment among low income blacks declined from the 1970s to 1990s such that whites are now more likely to enroll than blacks across the income spectrum (Black and Sufi 2002). In contrast to Black and Sufi's findings, this paper's analysis of the NLSY97 indicates that black youth are still more likely to enroll when demographics, income, school quality, and academic ability have been adjusted for.

The finding observed in models not adjusting for demographic factors that closely spaced siblings are more likely to complete college than widely spaced siblings is contrary to the resource dilution explanation, which holds that closely spaced youth will more heavily tax the available family resources than widely spaced youth. However, U.S. government financial aid formulas require families who have multiple children in college to pay less per child than equivalently sized families with only one child in college, which might allow closely spaced siblings to receive more financial support than they would have if they had been more widely spaced (US-Government 2013b). Alternatively, it is possible that families whose children have a high *a priori* probability of completing college had closely spaced siblings on purpose and the relationship is spurious.

In summary, this paper has examined potential explanations for the relationship between sibship size and educational attainment and variation in the relationship by race/ethnicity. It addressed the contention that the relationship is spurious, focusing on three aspects of families that could that could plausibly generate such a spurious relationship: parent religion, family structure, and parent education. It did not find support for the contention that the relationship is spurious, as the relationship remained significant when the three factors and many others had been adjusted for. We suspect that the sibship size penalty does exist and that it is causal even if its effects are relatively minor due to the compression of the family size

distribution over the late 20th century. If the relationship is not causal, the factors generating it are extremely subtle; they were neither directly measured by the many adjustment factors available in the NLSY97 nor indirectly measured via correlations with those factors. As with all social phenomena, "culture" could generate the relationship, but culture neither operates in isolation nor is monolithic. It varies by race, education, income, religion, and religiosity, all of which this analysis addressed. We welcome alternative explanations for the relationship between sibship size and educational attainment, but the results of this analysis imply that the relationship is both real and causal.

REFERENCES

Anastasi, Anne. 1956. "Intelligence and Family Size." *Psychological Bulletin* 53:187-209.

Bailey, Martha J., and Susan M. Dynarski. 2011. "Inequality in Post-Secondary Education." *Whither Opportunity*:117-32.

Becker, Gary S., James S. Duesenberry, and Bernard Okun. 1960. "An Economic Analysis of Fertility." Pp. 225-56 in *Demographic and Economic Change in Developed Countries*: Columbia University Press.

Becker, Gary S., and H. Gregg Lewis. 1973. "On the Interaction between Quantity and Quality of Children." *Journal of Political Economy* 82:279-87.

Black, Sandra, and Amir Sufi. 2002. "Who Goes to College? Differential Enrollment by Race and Family Background." in *NBER Working Paper Series*. Cambridge, MA: National Bureau of Economic Research.

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(1):84-94.

—. 1989. *Family Size and Achievement*. Los Angeles, CA: University of California Press.

Blau, Peter M., and Otis Dudley Duncan. 1967. *The American Occupational Structure*. New York: Wiley.

Booth, Alison L., and Hiau Joo Kee. 2008. "Birth order matters: the effect of family size and birth order on educational attainment." *Journal of Population Economics* 22(2):367-97.

Bras, Hilde, Jan Kok, and Kees Mandemakers. 2010. "Sibship Size and Status Attainment Across Contexts: Evidence from the Netherlands, 1840-1925." *Demographic Research* 23:73-104.

Brooks-Gunn, Jeanne, and Greg J. Duncan. 1997. "The Effects of Poverty on Children." *The Future of Children* 7(2):55-71.

Brooks-Gunn, Jeanne, and Lisa Markman. 2005. "The Contribution of Parenting to Ethnic and Racial Gaps in School Readiness." *The Future of Children* 15(1):139-68.

Bumpass, Larry, and Hsien-Hen Lu. 2000. "Trends in cohabitation and implications for children's family contexts in the United States." *Population Studies: A Journal of Demography* 54(1):29-41.

Cameron, Stephen V., and James J. Heckman. 2001. "The Dynamics of Educational Attainment for Black, Hispanic, and White Males." *Journal of Political Economy* 109(3):455-99.

Coleman, Marilyn, Lawrence Ganong, and Mark Fine. 2000. "Reinvestigating Remarriage: Another Decade of Progress." *Journal of Marriage and Family* 62(4):1288-307.

Cunha, Flavio, and James J. Heckman. 2007. "The Technology of Skill Formation." *American Economic Review* 97(2):31-47.

de Haan, Monique. 2010. "Birth order, family size and educational attainment." *Economics of Education Review* 29(4):576-88.

Downey, Douglas B. 1995. "When Bigger Is Not Better: Family Size, Parental Resources, and Children's Educational Performance." *American Sociological Review* 60(5):746-61.

—. 2001. "Number of Sibling and Intellectual Development: The Resource Dilution Explanation." *American Psychologist* 56:497-504.

Dumont, Arsene. 1890. *Dépopulation et civilisation: étude démographique.* Paris: Economica.

Featherman, David, and Robert M. Hauser. 1978. *Opportunity and Change*. New York: Academic Press.

Fomby, Paula, and Andrew J. Cherlin. 2007. "Family Instability and Child Well-Being." *American Sociological Review* 72(2):181-204.

Guo, Guang, and Leah K. VanWey. 1999. "Sibship Size and Intellectual Development: Is the Relationship Causal?" *American Sociological Review* 64(2):169-87.

Hauser, Robert M. 1993. "Trends in College Enrty among Whites, Blacks, and Hispanics." Pp. 61-104 in *Studies of Supply and Demand in Higher Education: A National Bureau of Economic Research Project Report*, edited by Charles T. Clotfelter and Michael Rothschild. Chicago: University of Chicago Press.

Heckman, James J., and Paul A. LaFontaine. 2010. "The American High School Graduation Rate: Trends and Levels." *The Review of Economics and Statistics* 99(2):244-62.

Hetherington, E. Mavis. 1992. "Coping with Marital Transitions: A Family Systems Perspective." *Monographs of the Society for Research in Child Development* 57(2-3):1-14.

Jaeger, M. M. 2008. "Do Large Sibships Really Lead to Lower Educational Attainment?: New Evidence from Quasi-Experimental Variation in Couples' Reproductive Capacity." *Acta Sociologica* 51(3):217-35.

Lareau, Annette. 2011. *Unequal childhoods: Class, race, and family life*. Berkeley, CA: University of California Press.

Mercy, James A., and Lala Carr Steelman. 1982. "Familial Influence on the Intellectual Attainment of Children." *American Sociological Review* 47:532-42.

National Longitudinal Surveys, The. 2013. "Aptitude, Achievement & Intelligence Scores."

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

Reardon, Sean F. 2011. "The Widening Academic Achievement Gap Between the Rich and the Poor: New Evidence and Possible Explanations." *Whither Opportunity*:91-116.

Steelman, Lala Carr. 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 Carr, and James A. Mercy. 1983. "Sex Differences in the Impact of the Number of Older and Number of Younger Siblings on IQ Performance." *Social Psychology Quarterly* 46:157-62.

Steelman, Lala Carr, Brian Powell, Regina Werum, and Scott Carter. 2002. "Reconsidering the Effects of Sibling Configuration: Recent Advances and Challenges." *Annual Review of Sociology* 28(1):243-69.

US-Census. 2012. "Current Population Survey - 2012 Annual Social and Economic Supplement." edited by U.S. Census Bureau. Washington, D.C.: U.S. Government Printing Office.

US-Government. 2013a. "America's Children: Key National Indicators of Well-Being, 2013." edited by Federal Interagency Forum on Child and Family Statistics. Washington, DC: U.S. Government Printing Offic.

—. 2013b. "The EFC Formula, 2013-2014." edited by Department of Education.

Wetzel, James R. 1990. "American Families: 75 Years of Change." *Monthly Labor Review* 113(3):4-13.

Wu, Lawrence L., and Brian C. Martinson. 1993. "Family Structure and the Risk of a Premarital Birth." *American Sociological Review* 58(2):210-32.

Zajonc, Robert B. 1976. "Family Configuration and Intelligence." *Science* 192:227-36.

Zajonc, Robert B., and Gregory B. Markus. 1975. "Birth order and Intellectual Development." *Psychological Review* 82:74-88.

CHAPTER 3

FAMILY STRUCTURE AND HIGHER EDUCATIONAL ATTAINMENT

Introduction

Families have changed since the mid-20th century. Divorce rates more than doubled between 1960 and 1980, then fell by a third between 1980 and 2006 (Amato 2010; US-Census 2008). By the year 2000, approximately half of marriages would eventually end in divorce (Schoen and Canudas-Romo 2006). In addition, births outside of marriage have increased precipitously, with rates climbing from 18% in 1980 to 39% in 2006 (Brown 2010; Martin et al. 2009). These changes have occurred in tandem with a significant increase in women's labor force participation, altering the economic dynamics of family life and the division of labor within two-parent households (Bianchi 2000).

Alongside debates about government policies promoting marriage, particularly in low-income populations, scholars have shown considerable interest in disentangling the effects that marriage, divorce, remarriage, single parenthood, and cohabitation have on the lives of children (Crosnoe and Cavanagh 2010; Graefe and Lichter 1999; Lichter, Graefe and Brown 2003; Nock 2005). Their studies have examined a host of child outcomes with a litany of methods, provoking significant debate about how family life should be measured, what outcomes should be of interest, and what methods are most appropriate for measuring the effects of families on the lives of children. Recognizing that the frequently observed association between family changes and negative child outcomes may be due to selection (i.e., parents with certain characteristics are both likely to experience family changes and to raise children having below average cognitive and non-cognitive skills), recent work has focused on identifying the portion of the association that is causal (Fomby and Cherlin 2007;

Gennetian 2005). This paper contributes to that literature by applying statistical techniques that adjust for selection to nationally representative data in order to estimate the causal effect of family change on youth enrollment in and completion of 4-year college.

Background

Beginning with the Moynihan Report's identification of a "tangle of pathology" in 1965, policy-makers and scholars have shown significant interest in the contribution of families to the life outcomes of youth and the production of inequality (McLanahan and Percheski 2008; Moynihan 1965). Because of the Moynihan Report's focus, much research concentrated on the role of single-mother families; however, a rich literature has emerged that examines the myriad family forms and processes that comprise the demographics of American families. As with the Moynihan Report, this literature is motivated by basic observations of the decreased cognitive skills, non-cognitive skills, and academic achievement of youth who live in non-traditional families (i.e., families without both biological parents present and/or with step or half siblings) (Biblarz and Raftery 1999; Cavanagh and Huston 2006; Ginther and Pollak 2004; Manning and Lamb 2003; McLanahan and Sandefur 1994).

Scholars have identified a number of mechanisms through which family structure might affect youth outcomes, including economic resources, parenting processes/resources, and adolescent adjustment (Brown 2010). Especially with the rise of two-earner households in the second half of the 20th century, single parent households due to divorce or non-marital births are likely to have significantly less income than traditional families (Manning and Brown 2006). This lack of income may lead to children being raised in poverty, which has been demonstrated to have negative effects on their life chances, or parental stress due to financial strain (Brooks-Gunn

and Duncan 1997; Demo and Fine 2010). Financial changes due to family disruption may also directly impact the way youth pay for college. The amount of family income and wealth has been shown to be related to parental contributions to college; because divorce typically decreases the pool of available resources, it is likely that college contributions will decrease as well (Conley 2001; Steelman and Powell 1991). Research by Turley and Desmond (2011) indicates that divorced parents contribute less to their children's college financing than do married biological parents. It also reports that remarried parents contribute a similar amount as divorced parents, even though their incomes are more similar to married biological parents. Additional research using different data confirm the finding that married biological parents provide more financial support for children in college than parents in other family structures (Henretta et al. 2012). Low levels of parental contributions to college financing lead to increased student borrowing and an increased likelihood that students will work while in college, which has been shown to decrease the likelihood of graduation (Bozick 2007).

Good parents provide guidance and support for their children; however, a prerequisite for the provision of these is time and emotional resources. It may be difficult for single parents to carve out enough time for optimal parenting since they must provide the household's economic resources and manage the minutiae of family life without the assistance of a partner. Parents experiencing marital turbulence or divorce may be emotionally drained, leading to less engaged parenting than they might otherwise provide (Sun 2001). Additionally, divorce is associated with non-resident fatherhood, which may present challenges for even the most dedicated fathers.

A theoretical linkage between family structure and youth outcomes that has received recent attention, and that this paper employs, is that of adolescent adjustment to change (Cavanagh and Huston 2006; Fomby and Cherlin 2007; Wu and Martinson

1993). This perspective hinges on the idea that family instability is harmful to children, regardless of the qualitative type of family transition that is occurring. For example, in the wake of a divorce children must adjust to new family routines due to the separation of parents and various other outcomes resulting from the impact of economic changes, such as new homes or new schools. When new persons enter the family due to remarriage, children must both acclimate to the new members' presence and negotiate the ways in which a new adult will parent the child. Hetherington and Kelly (2002) found that an adjustment period of up to 7 years occurs when a step-parent enters a family. Recent research has documented that family instability leads to a number of negative outcomes for youth including poorer performance in school, which may lead to decreased enrollment in and completion of college (Cavanagh and Fomby 2012; Cavanagh, Schiller and Riegle-Crumb 2006; Heard 2007). Recent research on the relationship between family instability and college enrollment and completion indicates that there is a negative relationship between the two, particularly when instability occurs early in a child's life (Fomby 2013). However, it is unclear how much of this relationship is causal in nature.

Much of the research on the effects of family structure and instability has failed to address the possibility that it is the characteristics of parents that lead both to family instability and poor academic performance of children (i.e., the relationship is due to "selection"). The difficulty of identifying causal effects stems from what has been called the "fundamental problem of causal inference": given some binary cause (such as experiencing a parental divorce), individuals can only be observed in *one* of the two possible causal states (experienced divorce or did not experience divorce) (Holland 1986; Morgan and Winship 2007). However, the "causal effect" of a given binary cause for an individual is defined as the difference on the outcome of interest when the individual did and did not experience the cause. For example, the causal

effect of a family divorce on a child's graduation from high school is the probability that the child would graduate with an intact family minus the probability that the same child would graduate with divorced parents. Clearly, reality only reveals to us one of these possibilities (the unobserved one is "counterfactual"), so we are always unable to estimate the causal effect for a given individual. Methods that involve random assignment of causes, such as experiments, circumvent this problem by assuming that the treatment and control groups are on average equivalent due to random assignment, so their group averages can serve as proxies for the average of individual-level causal effects we would ideally estimate. Unfortunately (or fortunately, depending on one's perspective), most aspects of family life cannot be subjected to random assignment due to ethical and practical issues. Therefore, researchers must rely on research designs that involve some sort of randomization not induced by a researcher (such as "natural experiments"), research designs that eliminate external influence by only making comparisons within subsets of units (such as fixed effect estimators), or statistical techniques that address selection by specifying a causal model linking cause to effect and adjusting for the causal antecedents that make treatment and control groups differ (Cook and Campbell 1979; Morgan and Winship 2007).

Fortunately, attention has recently been paid to the issue of causality in the family structure literature, and a number of studies have made serious attempts to estimate the portion of the observed relationships between families and child outcomes that is causal (Crosnoe and Cavanagh 2010). One approach has been to used fixed effects estimators to compare children within the same family who have different family histories (such as half-siblings) or to compare children to themselves by observing them before and after a family transition (Gennetian 2005; Li 2007). Another approach has been to use natural experiments such as the death of a parent (Biblarz and Gottainer 2000). Other scholars have used statistical techniques explicitly

designed to adjust for selection effects (Fomby and Cherlin 2007; Frisco, Muller and Frank 2007). While the results of these studies have been mixed, most have demonstrated that the unadjusted differences in outcomes between youth from traditional and other families shrink considerably once selection has been taken into account; many find that some portion of the relationship is causal. This paper applies a similar statistical technique to estimate the causal effect of family instability on enrollment in and completion of college while explicitly adjusting for selection effects.

Data

This examination of the relationship between family structure and educational attainment analyzes data from a large, nationally-representative survey conducted by the U.S. Bureau of Labor Statistics: the National Longitudinal Survey of Youth 1997 (NLSY97). This survey contains extensive and detailed data about each respondent's family, living situation, and educational experiences.

The NLSY97 was a probability sample of youth ages 12-16 in 1997 with additional surveys annually since 1997. The full sample is comprised of two sub-samples, cross-sectional and ethnic, both of which will be used for the analysis. Only respondents remaining in the sample after 13 waves are included in the analysis, yielding a final sample size of N=5,364.

The primary variable of interest for the analysis is family structure – a construct whose measurement presents conceptual challenges that will be discussed in detail below. The primary outcome variables of interest for this analysis are 4-year college enrollment and 4-year college completion. This paper does not analyze 2-year college attendance due to the historical U.S. policy focus on 4-year college attendance over 2-year college attendance and the additional analytical complexity added by

analyzing which youth attend which type of college. Of course, attending any form of higher education is encouraged by the U.S. government, whose programs serve both college types. However, the greatest returns to education are concentrated among 4-year college degrees, which are currently of primary importance with respect to higher educational inequality in the U.S. population.

All analyses use panel sample weights to make the data representative of the population from which the survey was drawn. Only respondents remaining in the sample after 13 waves were included in the estimates, but the weights are designed to maintain the sample's representativeness.

All item-missing data for model variables were singly imputed. For continuous variables, a regression model on the outcome was estimated among respondents with data, and the values for those missing data were imputed based on predicted values from the regression model. For binary variables, a logistic regression model with the variable as the outcome was estimated among those with data, and the values for respondents missing the variable were imputed based on the predicted probabilities from the model. For categorical variables, a hot-deck procedure was used to randomly select a variable value from sample members who matched the individual with missing data on a number of characteristics. The characteristics were race/ethnicity, whether their residence was in a rural area, whether they attended public or private school, whether they graduated from high school, whether they enrolled in a 4-year college, and whether they graduated from a 4-year college.

After presenting descriptive statistics and bivariate trends, the general strategy for the analysis is to estimate treatment effects for the different family structure histories using propensity-score models. The focus of the analysis is the causal effect of having a non-traditional family structure on enrollment in and completion of 4-year college.

Methods

Estimating causal effects requires postulating a model of the causal relationships linking the cause under investigation to the outcome of interest. Figure 1 displays a causal model linking family structure history to college enrollment and graduation. This model posits that the causal roots of a youth's family structure history begin with the characteristics of his parents' families. The parents' families and their experiences impact the characteristics of the parents: their levels of education, non-cognitive traits and behaviors, and health, among other things. In addition, the parents' families provide models of what relationships look like and the roles one should play in a family. Both of these things, the parents' characteristics and their beliefs about the roles of family members, determine whether they marry before they have children and whether they remain married throughout their children's childhoods, the components of family structure history. In turn, family structure directly impacts a child's intellectual and behavioral development, his academic performance, the family's income, and the quantity and type of actions a child takes to prepare for college attendance (such as taking the SAT/ACT, visiting potential colleges, etc.). These factors mediate the relationship between family structure and enrollment in college via a number of pathways, including school quality and plans to attend college. This model also posits that family structure directly affects the intellectual and behavioral development a youth undergoes while in college, which play a key role in whether or not the youth graduates. The key assumption required for the estimated causal effects to be valid is that there are no factors outside the model that cause family structure

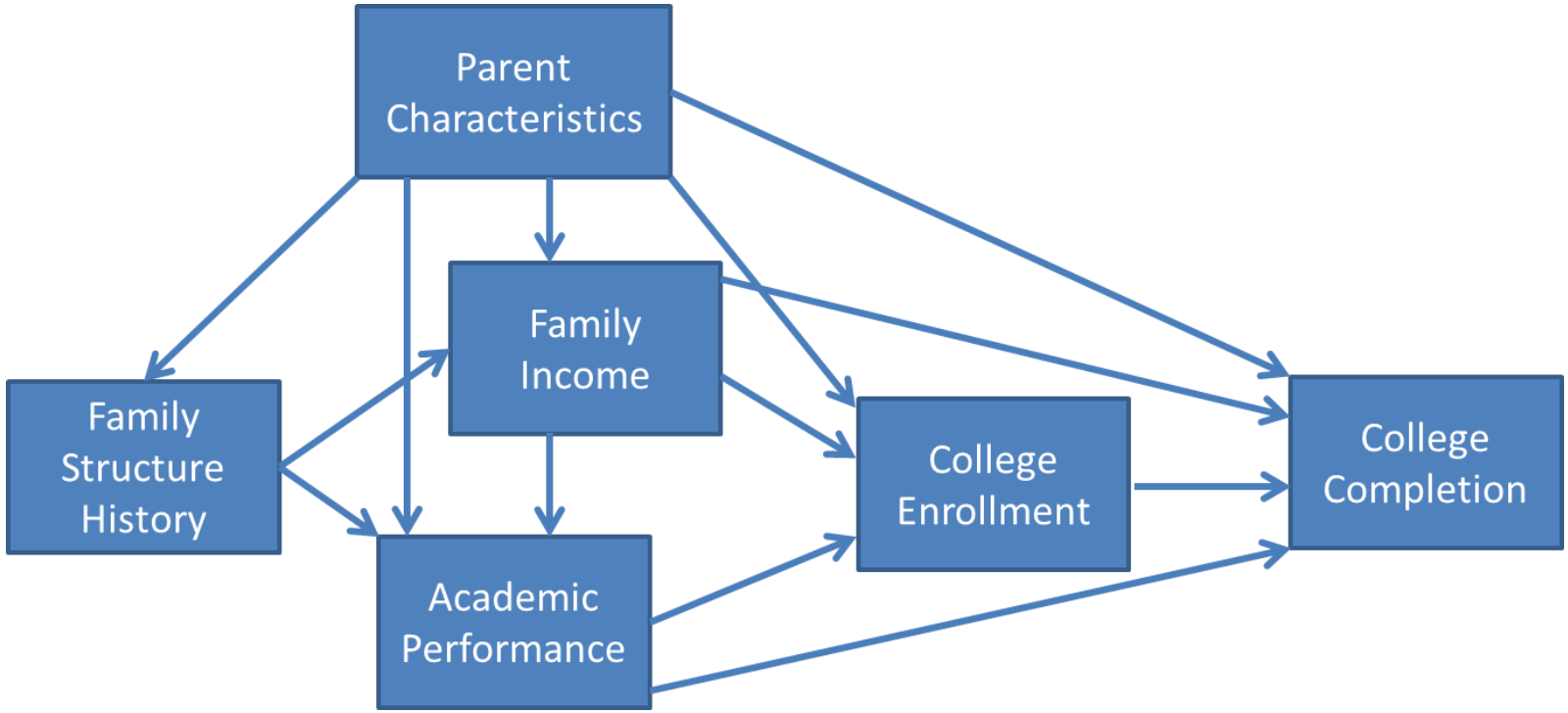


Figure 1 - Causal model linking family structure history to college enrollment and completion

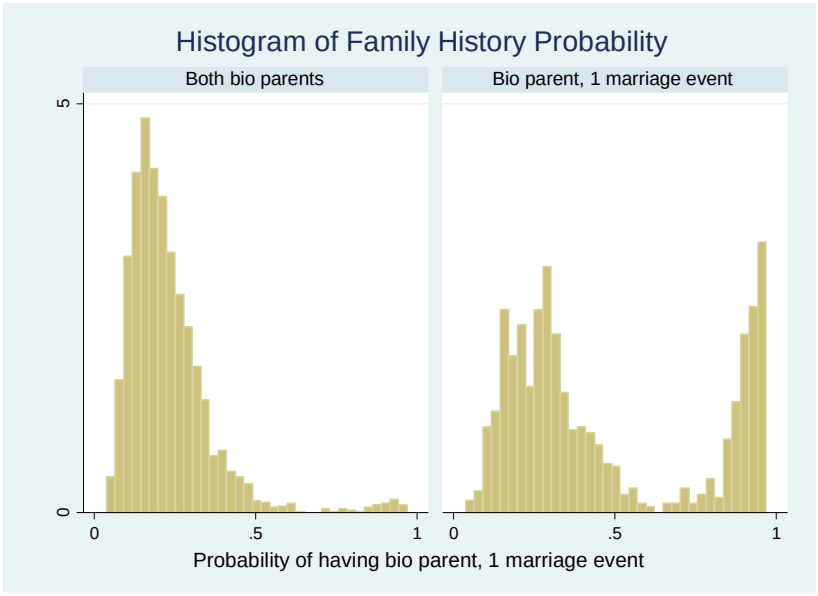
history or one of its antecedents and cause one of its descendants (i.e., there are no "backdoor paths" linking family structure history and enrollment in college).

Propensity score weight estimation of causal effects is composed of three phases: estimating the probability of receiving the treatment, constructing weights based on the estimated probability, and estimating the causal effect using the weights (Morgan and Winship 2007). For the paper's estimates, the "treatment" is experiencing a family history other than the traditional married biological parent family history. Each family history's effect on college enrollment is estimated relative to the traditional family, which is generally held to be the ideal for child development and economic security.

The probability of experiencing a non-traditional family structure is estimated via a series of logit models where the outcome is whether the youth lived in a traditional family structure or the alternative structure targeted by that model. These propensity models include predictors that are temporally prior to the youth's birth, gathered from the parent interview in the first wave of the survey. Model adjustment variables include the number of marriages and divorces the reporting parent experienced before the youth was born, the youth's race, whether the parent was an immigrant or the child of an immigrant, the region of the country they live in, whether they live in a rural area, the mother's age at her first birth, the mother's age at the youth's birth, and the educational attainment of both parents. Across the propensity models, marriage/divorce history, race, mother's age at youth's birth, and parental education tend to be the strongest predictors of family structure history. Not being married when the youth was born and having already experienced a marriage and a divorce are strongly related to experiencing a marriage transition after the youth is born. The older a mother is when she gave birth to a youth, the less likely she is to experience a marriage transition during the youth's childhood. Similarly, the more

education a youth's father received, the less likely the youth is to experience a marriage transition during his childhood. Somewhat counter intuitively, the more educated a youth's mother is, the more likely that youth is to experience a marriage transition during his childhood. This may reflect the economic dependence of relatively uneducated women on their partners, as women with higher earnings (or earning power) may be more willing to leave their husbands when displeased with a marriage.

The probability of experiencing one or more marriage transitions is estimated with moderate accuracy; model pseudo R squared values are approximately .2. In contrast, living with a single mother is predicted with very high accuracy, with a pseudo R squared value of .67. African-American youth are significantly more likely to live with single mothers for their entire childhood than are youth of other races. Living with a non-biological parent at some point is predicted with moderately high accuracy, with a pseudo R squared value of .35. Figure 2 displays the predicted probabilities of experiencing one marriage event and living with a single mother, as predicted by the propensity models.



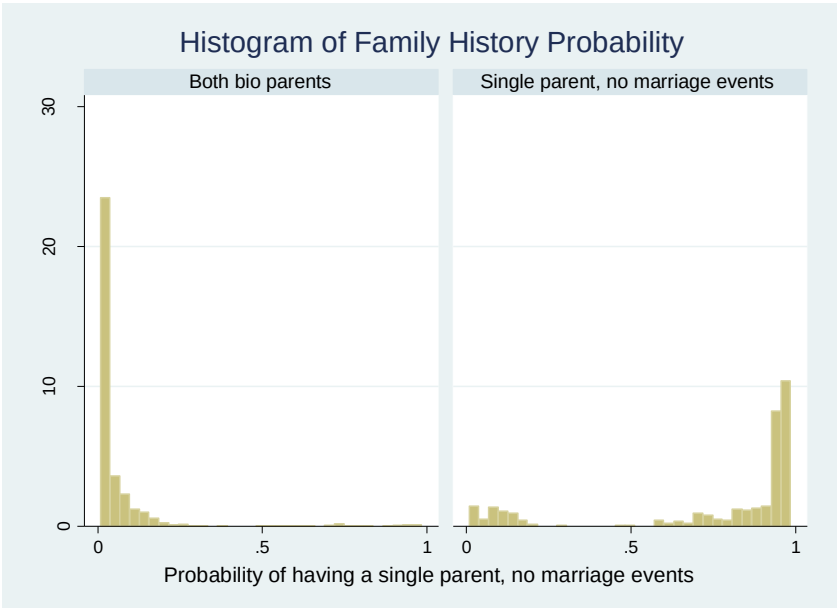


Figure 2 – Probability of experiencing one marriage event and living with a single parent, by observed outcome – National Longitudinal Survey of Youth 1997

Among the estimable parameters of a propensity score model are two of particular interest: the average treatment effect (ATE) and the average treatment effect among the treated (ATT). Conceptually, the average treatment effect estimates the treatment effect by propensity and averages over the predicted propensity distribution for the entire sample. In contrast, the average treatment effect among the treated averages the propensity-specific treatment effects with more weight given to those who experienced the treatment and those who did not experience the treatment but had a high predicted probability of experiencing it. The ATE estimates what the average effect of the treatment would be if people experienced the treatment with probability based on the entire predicted propensity distribution; the ATT estimates what the average effect of the treatment would be if people experienced the treatment in proportion with their predicted probability. In summary, the ATT is the average treatment effect among people who were likely to experience the treatment, and the ATE is the average treatment effect among everyone.

This paper proceeds by describing the population then presenting the ATE and ATT of family structure history on enrollment in and graduation from 4-year college.

Descriptive Analysis

Tables 1 and 2 display summary statistics for the survey. The U.S. is racially/ethnically diverse, with a significant number of Hispanic youth – although this analysis precedes the steady increase in Hispanic youth seen over the first decade of the 21st century. There are a significant number of immigrants present, with 15% of youth first or second-generation immigrants. Mothers are about 23 years old, on average, when they have their first child, and about 26 years old when they had the children aged 12-16 in 1997. Mothers and fathers attain approximately 13 years of

Table 1 - Summary Statistics - National Longitudinal Survey of Youth 1997

Variable	Values	Proportion	SE
Race/Ethnicity	White/Other	0.72	0.007
	Black	0.15	0.005
	Hispanic	0.13	0.005
Gender	Male	0.51	0.008
	Female	0.49	0.008
Immigrant Generation	1	0.05	0.003
	2	0.10	0.005
	3+	0.85	0.006
Region	Northeast	0.18	0.007
	North Central	0.28	0.008
	South	0.34	0.008
	West	0.20	0.007
Rural	Yes	0.28	0.008
	No	0.72	0.008
School type	Private/Parochial	0.10	0.005
	Public	0.90	0.005
Student-teacher ratio	<14	0.23	0.007
	14 to <18	0.40	0.008
	18 to <22	0.22	0.007
	22+	0.14	0.006
School size	<100	0.01	0.001
	100-299	0.07	0.004
	300-499	0.13	0.006
	500-749	0.24	0.007
	750 to 999	0.18	0.006
	1000+	0.38	0.008
Retained in grade	No	0.86	0.005
	Yes	0.14	0.005

education, on average. Household incomes are about $42,000, on average, in 1990 dollars. Most schools have student-teacher ratios of 14 to 18 students per teacher, most students attend school with 1000 students or more, and 14% of students have been retained in grade. Non-GED high school completion is about 83%, consistent with the

findings of Heckman and LaFontaine (2010). Approximately 48% of youth enroll in a 4-year college, with about 32% completing a 4-year degree by ages 25-29.

Measuring family structure requires care. Generally, family structure consists of who is present in a child's household and their relationships with one another – a married father and mother along with their children in a "traditional" family; a single mother (either divorced or never-married) in some families; and a biological parent and step-parent in other families. By definition, children have a family structure at all times from birth to age 17 (after which many leave home to live at college, with roommates, or on their own). Over these 17 years, many children will live in a household where family structure changes, perhaps due to divorce or remarriage. Therefore, a child whose family experiences a divorce has two family structures: he lived with both biological parents before the divorce and in some other configuration or configurations after.

Table 2 - Summary Statistics - National Longitudinal Survey of Youth 1997

Variable	Mean	SE
Graduate HS	0.83	0.006
Enroll 4-year College	0.48	0.008
Complete 4-year College	0.32	0.008
Age	14.68	0.021
Mother's age at R's birth	25.86	0.086
Mother's age at first birth	23.35	0.078
Mother's Highest Grade	13.12	0.046
Father's Highest Grade	12.93	0.05
Annual income (thousands of 1990 dollars)	42.19	0.596
ASVAB Paragraph Comprehension Scale Score	-166.01	14.12
ASVAB Math Scale Score	-258.19	14.851

How should this temporal variation be taken into account in measurement of family structure? If one measures family structure as the family configuration observed during a single measurement, one will over-simplify the child's family history (and may observe a family configuration that the child only experienced briefly). One could also define family structure in terms of whether a child's parents had undergone a marital transition or not. This definition would capture more information, but would not distinguish a child whose biological parents had divorced and remained unmarried from one whose parents were re-married (and perhaps brought other children into the marriage). One could also measure family structure as the configuration in which a child spent the most time; however, this definition would fail to capture the number of family structure transitions a child had experienced.

Family scholars have recognized this complexity in their examinations of family structure and have concluded that both the number and type of family transitions are important components of the effect of family structure on a child's life (Coleman, Ganong and Fine 2000; Fomby and Cherlin 2007; Wu and Martinson 1993). Each family transition requires a child to negotiate his or her relationships with the other, perhaps new, family members in the household. Some research has suggested that many children adjust in the wake of these transitions in two or three years (Hetherington 1992). This implies that most family structure transitions do not permanently affect children. However, it also implies that serial family transitions leave children in a state of continuous familial flux, which may impact their performance in school or their development more generally (depending on age). Table 3 displays summaries of these two characteristics, the qualitative and quantitative components of family transitions, for the population represented by the NLSY97.

Table 3 - Family History - National Longitudinal Survey of Youth 1997

Variable	Values	Proportion	SE
Family History	Both married bio parents since birth	0.48	0.01
	Bio parent since birth, marriages	0.35	0.01
	Single bio parent since birth, no marriages	0.06	0.00
	Non-bio parents at some point	0.11	0.00
Marital Changes	0	0.55	0.01
	1	0.20	0.01
	2	0.12	0.01
	3+	0.13	0.01

The first row of Table 3 summarizes the qualitative component of family history, indicating that about half of youth lived in a "traditional" family structure with both biological parents until age 17. Reflecting the divorce rates of the late 20th century, 35% of youth experienced one or more divorces or marriages during their childhood, along with potentially living with a single parent between marriages. In contrast, 6% of youth lived with a single biological parent who never married during their youth, although these youth could have experienced the presence of a cohabiting boyfriend or girlfriend of their parent. Finally, 11% of youth spent at least some time without a biological parent present in the household, which includes living with grandparents, other relatives, foster parents, or adoptive parents.

The second row of Table 3 summarizes the number of family transitions respondents of the NLSY97 experienced during their youth. Youth who lived with both biological parents since birth or with a single biological parent who never married experienced zero marital changes by definition. About 20% of youth experienced one marital transition, 12% experienced two transitions, and 13% experienced three or more transitions. This paper operationalizes family structure by

combining these two metrics, resulting in the family history summary presented in Table 4.

Table 4 - Combined Family History - National Longitudinal Survey of Youth 1997

	Proportion	**SE**
Both married bio parents since birth	0.48	0.01
Bio parent since birth, 1 marital change	0.18	0.01
Bio parent since birth, 2 marital changes	0.09	0.00
Bio parent since birth, 3+ marital changes	0.08	0.00
Single parent since birth	0.06	0.00
Non-bio parent at some point	0.11	0.00

The combined family history measure presented in Table 4 expands the family history measure from Table 3 by incorporating the number of marital changes experienced by youth who always lived with a biological parent but were present for one or more marital transitions. This measure indicates that 18% of youth always lived with a biological parent but experienced one marital transition (this category includes youth born into a family with both biological parents who then divorced and includes youth born to a single parent who then married), 9% of youth always lived with a biological parent but experienced two marital transitions (this category includes, for example, youth born into a family with both biological parents who then divorced and the parent the youth lived with re-married), and 8% of youth always lived with a biological parent but experienced three or more marital transitions. This combined family history measure captures both elements of family history and is the measure that will be used for the remainder of this paper.

Of note, cohabitation has become an important component of family formation and composition, increasing significantly between the late 20th century and early 21st century (Bumpass and Lu 2000). In 2012, 3.6% of children lived with one biological

parent and their cohabiting partner (US-Government 2013a). However, rates of cohabitation were lower in the 1980 to 2002 window during which NLSY97 respondents were 17 years old or younger. In the NLSY97 data, cohabiting partnerships were only observed after the survey began, whereas marriages were observed from the child's birth onwards. Including the observed cohabiting partnerships in the measurement of family structure history might bias inferences about cohabitation because only respondents whose parents cohabited while they were ages 12 to 17 would be identified, and they may not represent a random subset of respondents whose biological parents ever cohabited. Due to these measurement issues, only marriage relationships are included in the family structure history measure.

Table 5 presents the combined family history measure by race/ethnicity. Among families with white or other race youth, 54% lived with both biological parents since birth, a greater proportion than the 48% overall. 17% of youth experienced one marital change during their childhood, and 9% experienced two and three. Only three percent of white and other race youth lived with a single parent for

Table 5 - Combined Family History by Race/Ethnicity - National Longitudinal Survey of Youth 1997

	White/Other		Black		Hispanic	
	Prop.	SE	Prop.	SE	Prop.	SE
Both married bio parents since birth	0.54	0.01	0.18	0.01	0.48	0.02
Bio parent since birth, 1 marital change	0.17	0.01	0.21	0.01	0.19	0.01
Bio parent since birth, 2 marital changes	0.09	0.01	0.10	0.01	0.07	0.01
Bio parent since birth, 3+ marital changes	0.09	0.01	0.08	0.01	0.05	0.01
Single parent since birth	0.03	0.00	0.19	0.01	0.08	0.01
Non-bio parent at some point	0.08	0.01	0.24	0.01	0.13	0.01

their entire childhood, and 8% lived with non-biological parents at some point. In contrast to white and other race youth, 18% percent of black youth lived with both married biological parents from birth to age 17. Similar proportions of black youth experienced one, two, and three marital changes as white and other race youth. Significantly more black youth lived with a single parent throughout their childhood – 19%, compared to 3% of white and other race youth (although any of these single parents could have had cohabiting partners). A significantly greater proportion of black youth than white and other race youth lived without their biological parents at some point – 24%, compared to 8%. Finally, Hispanic youth have similar family histories to white and other race youth, with slightly lower rates of 2 and 3+ marital changes and slightly higher rates of living without biological parents at some point.

Table 6 presents the combined family history measure by income quartile. Income is probably affected by family structure, as having two adults present means that there could be two incomes contributing to household finances. Reflecting this fact, there is a smooth, monotonic increase in the proportion of youth living with both married biological parents as one moves up the household income scale, with 14% of youth in the lowest income quartile and 73% of youth in the highest income quartile living with both biological parents. Similarly, the proportion of youth experiencing one marital change monotonically decreases as one moves up the household income scale, with 28% of youth in the lowest income quartile and 8% of youth in the highest income quartile experiencing one marital change. The proportions of youth experiencing 2 and 3+ marital changes during their childhood are approximately equal across the income distribution, with a slightly lower proportion of youth in the highest income quartile experiencing 3+ marital changes.

Table 6 - Combined Family History by Household Income Quartile - National Longitudinal Survey of Youth 1997

	1 (Lowest)		2		3		4 (Highest)	
	Prop.	SE	Prop.	SE	Prop.	SE	Prop.	SE
Both married bio parents since birth	0.14	0.01	0.32	0.02	0.53	0.02	0.73	0.01
Bio parent since birth, 1 marital change	0.28	0.02	0.24	0.01	0.18	0.01	0.08	0.01
Bio parent since birth, 2 marital changes	0.09	0.01	0.08	0.01	0.09	0.01	0.08	0.01
Bio parent since birth, 3+ marital changes	0.11	0.01	0.12	0.01	0.09	0.01	0.05	0.01
Single parent since birth	0.19	0.01	0.08	0.01	0.02	0.00	0.01	0.00
Non-bio parent at some point	0.19	0.01	0.16	0.01	0.09	0.01	0.05	0.01

Reflecting the lower earning power of single-earner households, the proportion of youth who lived with a single non-married parent since birth significantly decreases as one moves up the income scale, with 19% of youth in the lowest income quartile, 8% in the second quartile, 2% in the third quartile, and only 1% in the fourth quartile. These rates of single parenting imply that it is heavily concentrated among those with lower incomes. Similarly, but with smaller decreases, youth with household incomes in the lowest two quartiles had higher rates of living without biological parents at some point than do youth in the upper two quartiles.

Table 7 presents rates of educational attainment by combined family history, including high school graduation, enrollment in 4-year college, and graduation from 4-year college. As one might anticipate from previous research, youth who had "traditional" married biological parent families attain education at rates above the overall average. In contrast, youth from all other family histories attain education at rates below the overall average.

Table 7 - 4-Year College Enrollment and Completion by Combined Family History - National Longitudinal Survey of Youth 1997

	HS Graduate		Enroll 4-Year College		Graduate 4-Year College	
	Proportion	**SE**	**Proportion**	**SE**	**Proportion**	**SE**
Overall	0.83	0.01	0.48	0.01	0.32	0.01
Both married bio parents since birth	0.92	0.01	0.63	0.01	0.46	0.01
Bio parent since birth, 1 marital change	0.79	0.01	0.42	0.02	0.25	0.02
Bio parent since birth, 2 marital changes	0.78	0.02	0.37	0.03	0.22	0.02
Bio parent since birth, 3+ marital changes	0.76	0.02	0.31	0.02	0.17	0.02
Single parent since birth	0.69	0.03	0.28	0.03	0.14	0.02
Non-bio parent at some point	0.69	0.02	0.27	0.02	0.13	0.02

High school graduation decreases from 92% among youth who lived with both married biological parents since birth to 69% among youth who lived with a single parent since birth or a non-biological parent at some point. Enrollment in 4-year college monotonically decreases as one moves from youth experiencing the stability of traditional families to the relative instability of families that experienced three or more marital transitions, with enrollment rates of 63% and 31%, respectively. 28% of youth from single parent families enroll in four-year college, and a similar proportion of youth enroll who lived with non-biological parents at some point. A similar pattern is present for graduation from 4-year college, with the highest proportion of graduating youth from traditional families and the lowest proportion of graduating youth living without a biological parent present at some point.

Table 8 presents two measures of family involvement in college financing: rates of families ever providing financial assistance for college and students ever receiving loans to pay for college among those who enrolled in 4-year college. Both of these variables have been shown to be related to 4-year completion rates, so they may play a mediating role between family history and college completion, especially via the relationship between family history and income due to the number of earners in a household (Bailey and Dynarski 2011; Dynarski 2000). 91% of youth who lived with

both biological parents since birth receive financial assistance from their family to help pay for college. The rate of providing family assistance drops about 10 percentage points among youth who experienced one or two marital changes, and another 10 percentage points to 71% among youth who experienced three or more marital changes, lived with a single biological parent since birth, or lived with a non-biological parent at some point. Assuming that the majority of parents who do not help their children pay for college do so due to financial constraints (and not due to preference), the economic impact of family dynamics is clearly reflected in rates of providing financial assistance during college.

Table 8 - College Financing by Combined Family History - National Longitudinal Survey of Youth 1997

	Family Help Pay		**Student Loans**	
	Proportion	**SE**	**Proportion**	**SE**
Overall	0.86	0.01	0.66	0.01
Both married bio parents since birth	0.91	0.01	0.62	0.01
Bio parent since birth, 1 marital change	0.82	0.02	0.71	0.03
Bio parent since birth, 2 marital changes	0.81	0.03	0.70	0.04
Bio parent since birth, 3+ marital changes	0.71	0.04	0.74	0.04
Single parent since birth	0.74	0.05	0.83	0.04
Non-bio parent at some point	0.71	0.04	0.65	0.04

Student loans are another important component of college financing, with two thirds of college students carrying student loan debt when they leave college. Student loans play a complementary role to that of family financial assistance; reflecting this, rates of student loan usage are lowest among youth who lived with both married biological parents since birth, at 62%. Student loan usage increases to about 71% among students who experienced 1 or more marital changes during their childhood. Student loan rates are highest, at 83%, among youth living with a single parent since

birth, reflecting the lower incomes among these families. Because college completion rates are lowest among these youth, they are most likely to have student loan debt without the increase in earnings associated with a college degree.

These results demonstrate that there is a relationship between family history and the likelihood of enrolling in and completing 4-year college – the remainder of this chapter estimates what proportion of this relationship is causal.

Propensity Models

As demonstrated by the rates of enrollment in 4-year college by family structure history and described in the causal model above, there is a non-trivial relationship between family structure history and rates of enrollment. The question, therefore, is how much of this relationship is causal. Do the familial and economic consequences of marital transitions or single parenthood disrupt the pathway to college, or are they simply correlated with the actual causal factors at play? This section estimates the magnitude of the causal effects of family structure history on enrollment in and completion of 4-year college.

Table 9 displays the estimated change in probability of enrolling in 4-year college due to having a given family structure history, relative to the probability of enrollment among youth from traditional families. As expected, the treatment effects are negative, reflecting the fact that youth from traditional families have the highest probability of enrolling in 4-year college. Experiencing one marital transition is associated with a significant drop of 10 percentage points in the probability of enrolling. As one might expect, experiencing more than one marital transition is associated with a larger decrease in the probability of enrolling, although the difference between experiencing one or more than one transition is not significant. Living in a single parent family throughout childhood is associated with a decrease in

the probability of enrollment of 18 percentage points, similar to that for experiencing 3 or more marital transitions. However, this effect is not significantly different than 0, which implies that the factors that lead to one living with a single parent are the same factors that ultimately determine enrollment in college, with family structure playing a minimal or non-causal role for youth from single parent families. This could also be due to the relatively small number of youth living with a single parent throughout their childhood, which decreases the model's ability to detect a causal effect. Living with a non-biological parent at some point has a similar effect on enrollment probability to that of experiencing 3 or more marital transitions, reflecting the family turbulence that probably leads to living without one's biological parents.

Table 9 - Combined Family History Treatment Effects on 4-Year College Enrollment - National Longitudinal Survey of Youth 1997

	Average Treatment Effect			**Average Treatment among the Treated**		
	Δ Probability	**SE**	**P-Value**	**Δ Probability**	**SE**	**P-Value**
Both bio parents since birth	Reference			Reference		
Bio parent since birth, 1 marital change	-0.10	0.028	<0.001	-0.08	0.035	0.021
Bio parent since birth, 2 marital changes	-0.12	0.043	0.004	-0.12	0.042	0.005
Bio parent since birth, 3+ marital changes	-0.22	0.055	<0.001	-0.14	0.041	0.001
Single parent since birth	-0.18	0.101	0.073	-0.07	0.083	0.370
Non-bio parent at some point	-0.22	0.056	<0.001	-0.11	0.053	0.033

Compared to the average treatment effects in the population, the average treatment effects among the treated are uniformly equal or smaller. This implies that the causal effects of non-traditional families are concentrated among those who are less likely to experience them, which will be examined in more detail below. However, with the exception of living with a single parent, the treatment effects for the treated are still all negative and significant. Using the overall population probability of enrolling of .48 as a reference point, the estimated causal effects of non-

traditional family structure histories are relatively large in magnitude: the ATT estimated decrease of .08 associated with experiencing one marital transition reflects a 17% percent drop in the likelihood of enrolling in college (.08/.48 = .17), certainly a non-trivial decrease.

Table 10 displays the estimated change in probability of graduating from 4-year college due to having a given family structure history, relative to the probability of graduation among youth from traditional families. The ATEs are similar to or larger in magnitude than the corresponding effects for enrollment, with decreases of about 11 percentage points in the probability of graduating for youth who experienced one or two marital changes. The negative effects for youth experiencing three marital

Table 10 - Combined Family History Treatment Effects on 4-Year College Completion - National Longitudinal Survey of Youth 1997

	Average Treatment Effect			**Average Treatment among the Treated**		
	Δ Probability	**SE**	**P-Value**	**Δ Probability**	**SE**	**P-Value**
Both bio parents since birth	Reference			Reference		
Bio parent since birth, 1 marital change	-0.11	0.03	<0.001	-0.06	0.03	0.022
Bio parent since birth, 2 marital changes	-0.11	0.05	0.018	-0.08	0.03	0.013
Bio parent since birth, 3+ marital changes	-0.26	0.04	<0.001	-0.10	0.03	0.001
Single parent since birth	-0.19	0.08	0.012	-0.04	0.07	0.516
Non-bio parent at some point	-0.16	0.06	0.006	-0.07	0.03	0.025

changes are significantly larger than those for one or two, a 26 percentage point decrease. In contrast with the enrollment estimates, the average treatment effect for living with a single parent since birth is measured with enough precision to attain significance. The single parent effect is between those for one or two marital changes and those for three marital changes, with a 19 percentage point decrease in the probability of completing 4-year college associated with growing up with a single

parent. Living with a non-biological parent at some point during childhood is associated with a 16 percentage point drop in the probability of completing college.

As with the enrollment estimates, the ATT effects are uniformly smaller in magnitude than the ATEs, often significantly. Most ATTs are about half of their respective ATEs, with the exception of that associated with being raised by a single parent, which decreases by a factor of about 5 and becomes insignificant.

These results imply that there are real consequences, both familial and financial, to departures from a traditional family structure. However, the results do not imply that divorce or single parenthood are uniformly negative; there are situations when all parties are better off with marriage dissolution. The results do mean, however, that on average children fare worse with respect to higher education when raised in a non-traditional family, and that there is a causal relationship between the two.

Discussion

Similar to previous research, this investigation has found that there is a non-negligible causal relationship between family instability and college enrollment and completion (Fomby 2013). However, it has also found that these effects differ depending on whether one examines the entire population or the portion of the population that is likely to experience unstable families.

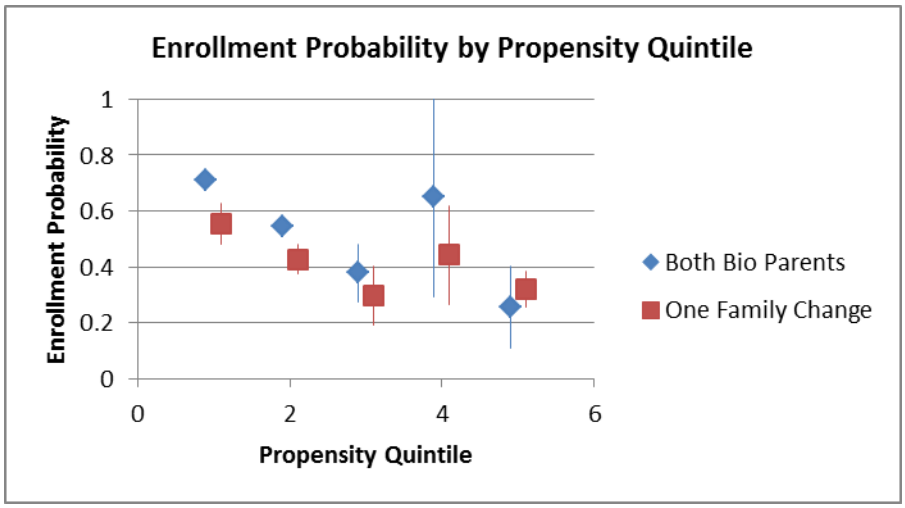
The models reveal that the estimated average treatment effects among the treated are smaller than the average treatment effects among the population: why? It is informative to examine the estimated treatment effects by estimated propensity of receiving the treatment, effectively unpacking what the weighted propensity models are estimating. Figure 3 displays the estimated probability of enrollment in 4-year college by propensity quintile and observed family structure history for one marital

transition and single parent histories, with 95% confidence intervals around the estimated probabilities.

The first thing of note about Figure 3 is that the width of the confidence intervals is a function of the distribution of estimated propensities shown in Figure 2. Because few youth who live in traditional families have a high propensity for experiencing another family history, the "both bio parents" confidence interval widths are largest in the higher propensity quintiles. Conversely, the confidence interval widths for youth in non-traditional families are smallest in the higher propensity quintiles. Additionally, there are not enough cases in the third quintile of propensity .4 - .6 for the single parent model to estimate enrollment probabilities.

For both graphs in Figure 3, the treatment effects are concentrated in the lowest propensity quintiles, with no significant effects observed in the highest two quintiles. Because the ATEs weight based on the entire propensity distribution, the lowest propensity quintiles contribute a significant amount to them; in contrast, the ATTs put more emphasis on the highest propensity quintiles where the treatment effects are small or non-existent.

The concentration of effects in the lowest quintiles is important because it contextualizes the average treatment effects estimated above by demonstrating that they are heterogeneous among the population (Morgan and Todd 2008). This implies two related facts about the effects of non-traditional family structures: first, those with the highest likelihood of experiencing non-traditional family structures are least



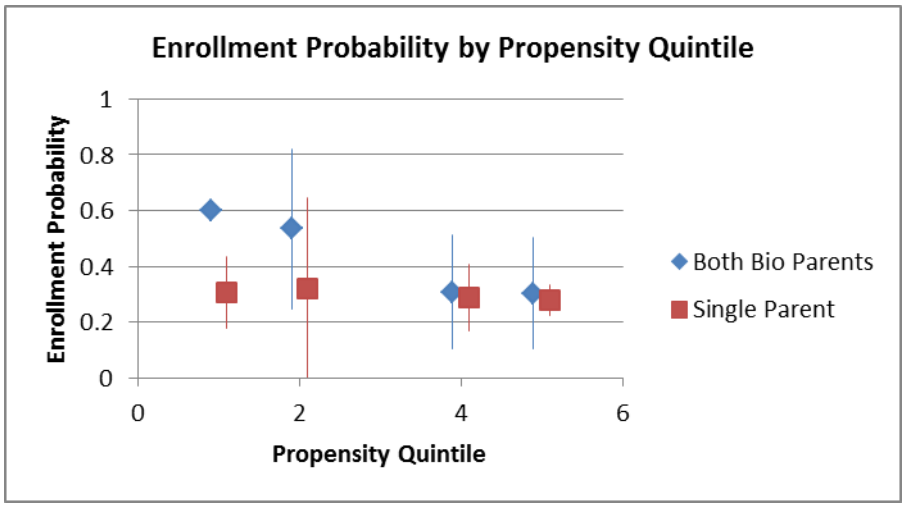


Figure 3 – Probability of enrollment in 4-year college by propensity quintile and observed family structure history – National Longitudinal Survey of Youth 1997

affected by them. This means that many youth who are raised in non-traditional families and do not enroll in college are doing so for reasons other than their family structure – perhaps due to the educational or neighborhood contexts in which they were raised, and the effects of these factors on their academic performance and actions to prepare for college. Second, and probably more importantly, those who are most likely to have traditional family structures, but do not, experience the brunt of the

negative causal effects of family structure on enrollment in 4-year college. This means that among the 18% of youth whose family undergoes one marital transition (Table 4), approximately 60% have a low probability of experiencing that transition and are therefore exposed to the negative causal effects of the family transition (Figure 2). This implies that about 10% of all youth are subject to the approximate 10 percentage point decrease in the likelihood of enrolling in 4-year college due to experiencing one family marriage transition. Repeating this exercise and summing across all non-traditional family types, approximately 25% of all youth are exposed to the negative causal effects of a non-traditional family structure history by having a low propensity to experience a non-traditional family structure but experiencing one anyway.

There are multiple potential explanations for the absence of treatment effects among those with a high propensity to experience a non-traditional family structure. One explanation involves which youth are likely to enroll in college in settings where few youth enroll. A simple model of college enrollment patterns might posit that youth are tacitly ranked on intelligence and work ethic during high school, and that the youth who are highly ranked enroll in college. Furthermore, some youth have been identified as promising from childhood so their high ranking has been assured by adult investment that less promising youth do not receive. Let us assume that the rankings are approximately normally distributed and that some constant proportion of youth in all communities, say 30%, receives adult investment and are therefore likely to enroll in college regardless of other contextual factors such as family structure and dynamics. In contrast, the other 70% of youth receive less adult investment, so their ranking has higher variance over time and their enrollment is more contingent on contextual factors such as family structure, peer behavior, or being assigned a particularly gifted or inspiring teacher. In communities where few youth enroll in college, the “bar” for ranking high enough to enroll in college is rather high, so most of the youth who enroll

are the 30% of youth who have received additional investment. Because the ranking variance over time of these youth is low and their enrollment is relatively nonresponsive to contextual factors, few youth who are on the enrollment margin are pushed across by their family structure. In contrast, in communities where many youth attend college, the ranking "bar" for college enrollment is low and many youth who enroll did not receive the additional adult investment. Because the ranking variance over time of youth near the mean of the ranking distribution is high and their enrollment is responsive to contextual factors, changes in family structure impact the youth on the enrollment margin enough to push them across the "bar" for enrollment. In addition to lower-ranked individuals possibly being more sensitive to contextual factors due to less adult investment in them, the shape of a normal distribution of rank guarantees that more individuals will be on the margins of a bar that is closer to the mean of the distribution. Thus, a simple threshold model of enrollment driven by low temporal variance among those ranked high and high temporal variance among those ranked middle/low and the basic facts of normal distributions could plausibly generate the observed treatment effect pattern.

A second, perhaps more sociological, explanation for the absence of treatment effects among those with a high propensity to experience a non-traditional family structure involve the communities in which they reside. This explanation posits that communities adapt to the dominant family structures of their members. In communities where non-traditional family structures are rare, the nuclear family is a reliable and stable vehicle for the development of children, so the surplus resources of the community are mobilized in ways that presuppose and complement the presence and functionality of traditional families. For example, surplus time is expended as parental volunteering at schools, and community funds contribute to collective goods such as block parties and neighborhood watch organizations that do not substitute for

family-provided goods. In these communities, having a stable family is presupposed, so no resources need be available for replacing or supplementing goods traditionally provided by families - tangible, emotional, or educational in nature. In these communities, enjoying communal goods requires that the more critical needs provided by families have been met. In contrast, community resources are expended in different ways in communities where traditional families are rare and propensities for non-traditional family structures are high (and stigma low). In these communities, limited surplus resources are expended to provide goods that in some ways substitute for the absence of the support provided by traditional families, such as watching each other's children, community centers that provide care and amusement for youth and older adults and classes on parenting and financial literacy. Additionally, the expected roles of extended family members such as grandparents reflect the absence of traditional families – they act as caregivers, both emotionally and financially, whereas they might spend their time in other ways if their children's families did not require their contributions (not to say that grandparents do not provide support in communities where most families are traditional, but residential mobility is probably higher in these communities such that fewer families are near the grandparents). As noted earlier, the essential resources of money and time are at play here: parents in non-traditional families tend to have less of each, so they rely on others outside the nuclear family. Of course, family structure is neither deterministic nor monolithic – there are single parents who provide better childhood environments than most couples, step-families whose presence enriches the lives of all children, communities of traditional families where inter-familial support networks are extensive and robust, and communities of non-traditional families where each family's needs are too great to provide for anyone else's. However, communities allocate resources in response to their members' needs

and values, and those needs and values are shaped by the types of families in which people find themselves.

A third explanation hinges on the relationship between income and college education (Bailey and Dynarski 2011). If one of the mechanisms through which family instability affects college enrollment and completion is income, it stands to reason that individuals with more income to lose via divorce will be more affected by it. In contrast, individuals whose incomes are relatively low will not experience significant losses due to divorce, and therefore will not experience the portion of family instability's effects that operates via income. Research indicates that low-income mothers experience relatively few economic benefits from marriage, so, conversely, they likely experience fewer economic detriments from divorce than do higher income individuals (Lichter, Qian and Mellott 2006).

As is usually the case, none of these proposed explanations for the concentration of family structure effects among those least likely to experience them is the complete story. All three explanations are likely to reflect reality to some degree - in some places more one than the others, and in other places all three in tandem. Nevertheless, these simple models provide plausible explanations for the heterogeneity of family structure effects observed in the NLSY97 data.

This paper has found that there are non-negligible and significant effects of family structure and instability on the likelihood that youth enroll in and complete 4-year college. It has also found that these effects are heterogeneous – the causal impact is negligible among youth who are likely to experience family instability, while they are significant among youth who are unlikely to experience family instability. Family instability increased significantly throughout the 20th century, and new family forms are emerging with increased rates of cohabitation; future research will determine whether these new family forms affect youth in the same way as divorce and

remarriage. With a college education becoming ever more critical to economic success and stability, school and government policies should recognize and respond to the financial and academic difficulties caused by family instability.

REFERENCES

Adelman, Clifford. 2006. "The Toolbox Revisited: Paths to Degree Completion From High School Through College." *US Department of Education*.

Amato, Paul R. 2010. "Research on Divorce: Continuing Trends and New Developments." *Journal of Marriage and Family* 72(3):650-66.

Anastasi, Anne. 1956. "Intelligence and Family Size." *Psychological Bulletin* 53:187-209.

Bailey, Martha J., and Susan M. Dynarski. 2011. "Inequality in Post-Secondary Education." *Whither Opportunity*:117-32.

Becker, Gary S., James S. Duesenberry, and Bernard Okun. 1960. "An Economic Analysis of Fertility." Pp. 225-56 in *Demographic and Economic Change in Developed Countries*: Columbia University Press.

Becker, Gary S., and H. Gregg Lewis. 1973. "On the Interaction between Quantity and Quality of Children." *Journal of Political Economy* 82:279-87.

Bianchi, S.M. 2000. "Maternal employment and time with children: Dramatic change or surprising continuity?" *Demography* 37(4):401-14.

Biblarz, T.J., and G. Gottainer. 2000. "Family structure and children's success: A comparison of widowed and divorced single-mother families." *Journal of Marriage and Family* 62(2):533-48.

Biblarz, T.J., and A.E. Raftery. 1999. "Family Structure, Educational Attainment, and Socioeconomic Success: Rethinking the 'Pathology of Matriarchy'." *American Journal of Sociology* 105(2):321-65.

Black, Sandra, and Amir Sufi. 2002. "Who Goes to College? Differential Enrollment by Race and Family Background." in *NBER Working Paper Series*. Cambridge, MA: National Bureau of Economic Research.

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(1):84-94.

—. 1989. *Family Size and Achievement*. Los Angeles, CA: University of California Press.

Blau, Peter M., and Otis Dudley Duncan. 1967. *The American Occupational Structure*. New York: Wiley.

Booth, Alison L., and Hiau Joo Kee. 2008. "Birth order matters: the effect of family size and birth order on educational attainment." *Journal of Population Economics* 22(2):367-97.

Bozick, Robert. 2007. "Making It Through the First Year of College: The Role of Students' Economic Resources, Employment, and Living Arrangement." *Sociology of Education* 80(3):261-85.

Bras, Hilde, Jan Kok, and Kees Mandemakers. 2010. "Sibship Size and Status Attainment Across Contexts: Evidence from the Netherlands, 1840-1925." *Demographic Research* 23:73-104.

Brooks-Gunn, Jeanne, and Greg J. Duncan. 1997. "The Effects of Poverty on Children." *The Future of Children* 7(2):55-71.
Brooks-Gunn, Jeanne, and Lisa Markman. 2005. "The Contribution of Parenting to Ethnic and Racial Gaps in School Readiness." *The Future of Children* 15(1):139-68.
Brown, Susan L. 2010. "Marriage and Child Well-Being: Research and Policy Perspectives." *Journal of Marriage and Family* 72(5):1059-77.
Bumpass, Larry, and Hsien-Hen Lu. 2000. "Trends in cohabitation and implications for children's family contexts in the United States." *Population Studies: A Journal of Demography* 54(1):29-41.
Cameron, Stephen V., and James J. Heckman. 2001. "The Dynamics of Educational Attainment for Black, Hispanic, and White Males." *Journal of Political Economy* 109(3):455-99.
Card, David E. 1999. "The Causal Effect of Education on Earnings." *Handbook of Labor Economics* 3:1801-63.
Cavanagh, S.E., and P. Fomby. 2012. "School context, family instability, and the academic careers of adolescents." *Sociology of Education* 85(1):81-97.
Cavanagh, S.E., and A.C. Huston. 2006. "Family instability and children's early problem behavior." *Social Forces* 85(1):551-81.
Cavanagh, S.E., K.S. Schiller, and C. Riegle-Crumb. 2006. "Marital transitions, parenting, and schooling: Exploring the link between family-structure history and adolescents' academic status." *Sociology of Education* 79(4):329-54.
Coleman, Marilyn, Lawrence Ganong, and Mark Fine. 2000. "Reinvestigating Remarriage: Another Decade of Progress." *Journal of Marriage and Family* 62(4):1288-307.
College Board, The. 1998. "Trends in Student Aid." Washington, D.C.
—. 2006. "Trend in Student Aid." Washington, D.C.
Conley, Dalton. 2001. "Capital for College: Parental Assets and Postsecondary Schooling." *Sociology of Education* 74(1):59-72.
Cook, Thomas D., and Donald T. Campbell. 1979. *Quasi-Experimentation: Design & Analysis Issues for Field Settings*. Chicago: Rand McNally.
Crosnoe, Robert, and Shannon E. Cavanagh. 2010. "Families With Children and Adolescents: A Review, Critique, and Future Agenda." *Journal of Marriage and Family* 72(3):594-611.
Cunha, Flavio, and James J. Heckman. 2007. "The Technology of Skill Formation." *American Economic Review* 97(2):31-47.
de Haan, Monique. 2010. "Birth order, family size and educational attainment." *Economics of Education Review* 29(4):576-88.
Demo, D.H., and M.A. Fine. 2010. *Beyond the average divorce*. Thousand Oaks, CA: Sage.
Downey, Douglas B. 1995. "When Bigger Is Not Better: Family Size, Parental Resources, and Children's Educational Performance." *American Sociological Review* 60(5):746-61.
—. 2001. "Number of Sibling and Intellectual Development: The Resource Dilution Explanation." *American Psychologist* 56:497-504.

Dumont, Arsene. 1890. *Dépopulation et civilisation: étude démographique.* Paris: Economica.
Dynarski, Susan M. 2000. "Hope for Whom? Financial Aid for the Middle Class and Its Impact on College Attendance." *National Tax Journal* 53(3):629-61.
Dynarski, Susan M., and Judith E. Scott-Clayton. 2006. "The Cost of Complexity in Federal Student Aid: Lessons from Optimal Tax Theory and Behavioral Economics." *National Tax Journal* 59(2):319-56.
Featherman, David, and Robert M. Hauser. 1978. *Opportunity and Change*. New York: Academic Press.
Fomby, Paula. 2013. "Family Instability and College Enrollment and Completion." *Population Research and Policy Review* 32(4):469-94.
Fomby, Paula, and Andrew J. Cherlin. 2007. "Family Instability and Child Well-Being." *American Sociological Review* 72(2):181-204.
Frisco, Michelle L., Chandra Muller, and Kenneth Frank. 2007. "Parents' Union Dissolution and Adolescents' School Performance: Comparing Methodological Approaches." *Journal of Marriage and Family* 69(3):721-41.
Gennetian, Lisa A. 2005. "One or two parents? Half or step siblings? The effect of family structure on young children's achievement." *Journal of Population Economics* 18(3):415-36.
Ginther, Donna K., and Robert A. Pollak. 2004. "Family Structure and Children's Educational Outcomes: Blended Families, Stylized Facts, and Descriptive Regressions." *Demography* 41(4):671-96.
Goldin, Claudia Dale, and Lawrence F. Katz. 2008. *The Race Between Education and Technology*. Cambridge, Mass.: Belknap Press of Harvard University Press.
Graefe, Deborah Roemke, and Daniel T. Lichter. 1999. "Life Course Transitions of American Children: Parental Cohabitation, Marriage, and Single Motherhood." *Demography* 36(2):205-17.
Guo, Guang, and Leah K. VanWey. 1999. "Sibship Size and Intellectual Development: Is the Relationship Causal?" *American Sociological Review* 64(2):169-87.
Hauser, Robert M. 1993. "Trends in College Enrty among Whites, Blacks, and Hispanics." Pp. 61-104 in *Studies of Supply and Demand in Higher Education: A National Bureau of Economic Research Project Report*, edited by Charles T. Clotfelter and Michael Rothschild. Chicago: University of Chicago Press.
Heard, H.E. 2007. "The family structure trajectory and adolescent school performance." *Journal of Family Issues* 28(3):319-54.
Heckman, James J., and Paul A. LaFontaine. 2010. "The American High School Graduation Rate: Trends and Levels." *The Review of Economics and Statistics* 99(2):244-62.
Henretta, John C., Douglas A. Wolf, Matthew F. Van Voorhis, and Beth J. Soldo. 2012. "Family structure and the reproduction of inequality: Parents' contribution to children's college costs." *Social Science Research* 41(4):876-87.

Hetherington, E. Mavis. 1992. "Coping with Marital Transitions: A Family Systems Perspective." *Monographs of the Society for Research in Child Development* 57(2-3):1-14.
Hetherington, E.M., and J. Kelly. 2002. *For better or for worse: divorce reconsidered*. New York: W.W. Norton.
Holland, Paul W. 1986. "Statistics and Causal Inference." *Journal of the American Statistical Association* 81(396):945-70.
Jaeger, M. M. 2008. "Do Large Sibships Really Lead to Lower Educational Attainment?: New Evidence from Quasi-Experimental Variation in Couples' Reproductive Capacity." *Acta Sociologica* 51(3):217-35.
Lareau, Annette. 2011. *Unequal childhoods: Class, race, and family life*. Berkeley, CA: University of California Press.
Li, J.C.A. 2007. "The kids are OK: Divorce and children's behavior problems." in *RAND Labor and Population Working Paper Series No. WR-489.* Santa Monica, CA: Rand.
Lichter, D.T., Z. Qian, and L.M. Mellott. 2006. "Marriage or dissolution? Union transitions among poor cohabiting women." *Demography* 43(2):223-40.
Lichter, Daniel T., Deborah Roemke Graefe, and J. Brian Brown. 2003. "Is Marriage a Panacea? Union Formation Among Economically Disadvantaged Unwed Mothers." *Social Problems* 50(1):60-86.
Long, Bridget T. 2004. "The Impact of Federal Tax Credits for Higher Education Expenses." in *College Choices: The Economics of Where to Go, When to Go, and How to Pay For It*, edited by Caroline M. Hoxby. Chicago, IL: University of Chicago Press.
Manning, W.D., and S.L. Brown. 2006. "Children's economic well-being in marriage and cohabiting parent families." *Journal of Marriage and Family* 68(2):345-62.
Manning, W.D., and K.A. Lamb. 2003. "Adolescent well-being in cohabiting, married, and single-parent families." *Journal of Marriage and Family* 65(4):876-93.
Martin, J.A., B.E. Hamilton, P.D. Sutton, S.J. Ventura, F. Menacker, S. Kirmeyer, and T.J. Matthews. 2009. "Births: Final data for 2006 - National Vistal Statistics Reports." Hyattsville, MD: National Center for Health Statistics.
McLanahan, Sara S., and Christine Percheski. 2008. "Family Structure and the Reproduction of Inequalities." *Annual Review of Sociology* 34:257-76.
McLanahan, Sara S., and G. Sandefur. 1994. *Growing Up With a Single Parent: What Hurts, What Helps*. Cambridge, MA: Harvard University Press.
Mercy, James A., and Lala Carr Steelman. 1982. "Familial Influence on the Intellectual Attainment of Children." *American Sociological Review* 47:532-42.
Morgan, Stephen L., and Jennifer J. Todd. 2008. "A Diagnostic Routine for the Detection of Consequential Heterogeneity of Causal Effects." *Sociological Methodology* 38(1):231-81.

Morgan, Stephen L., and Christopher Winship. 2007. *Counterfactuals and Causal Inference: Methods and Principles for Social Research*. New York: Cambridge University Press.
Moynihan, Daniel P. 1965. "The Negro Family: The Case for National Action." edited by Office of Policy Planning and Research. Washington, DC: U.S. Department of Labor.
National Longitudinal Surveys, The. 2013a. "Aptitude, Achievement & Intelligence Scores."
—. 2013b. "Constructing Comparable Samples across the NLSY79 and NLS&97."
Nock, Steven L. 2005. "Marriage as a Public Issue." *The Future of Children* 15(2):13-32.
Powell, Brian, and Lala Carr Steelman. 1990. "Beyond Sibship Size: Sibling Density, Sex Composition, and Educational Outcomes." *Social Forces* 69(1):181-206.
Reardon, Sean F. 2011. "The Widening Academic Achievement Gap Between the Rich and the Poor: New Evidence and Possible Explanations." *Whither Opportunity*:91-116.
Schoen, R, and V Canudas-Romo. 2006. "Timing effects on divorce: 20th century experience in the United States." *Journal of Marriage and Family* 68(3):749-58.
Steelman, Lala Carr. 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 Carr, and James A. Mercy. 1983. "Sex Differences in the Impact of the Number of Older and Number of Younger Siblings on IQ Performance." *Social Psychology Quarterly* 46:157-62.
Steelman, Lala Carr, Brian Powell, Regina Werum, and Scott Carter. 2002. "Reconsidering the Effects of Sibling Configuration: Recent Advances and Challenges." *Annual Review of Sociology* 28(1):243-69.
Steelman, Lala Carr, and Brian; Steelman Powell, Lala Carr. 1991. "Sponsoring the Next Generation: Parental Willingness to Pay for Higher Education." *American Journal of Sociology* 96(6):1505-29.
Sun, Y. 2001. "Family environment and adolescents' well-being before and after parents' marital disruption: A longitudinal analysis." *Journal of Marriage and Family* 63(3):697-713.
Turley, Ruth N. Lopez, and Matthew Desmond. 2011. "Contributions to College Costs by Married, Divorced, and Remarried Parents." *Journal of Family Issues* 32(6):767-90.
US-Census. 2008. "Statistical abstract of the United States: 2008." edited by U.S. Census Bureau. Washington, DC: U.S. Government Printing Office.
—. 2012. "Current Population Survey - 2012 Annual Social and Economic Supplement." edited by U.S. Census Bureau. Washington, D.C.: U.S. Government Printing Office.
US-Government. 2013a. "America's Children: Key National Indicators of Well-Being, 2013." edited by Federal Interagency Forum on Child and Family Statistics. Washington, DC: U.S. Government Printing Offic.

—. 2013b. "The EFC Formula, 2013-2014." edited by Department of Education.
Wetzel, James R. 1990. "American Families: 75 Years of Change." *Monthly Labor Review* 113(3):4-13.
Wu, Lawrence L., and Brian C. Martinson. 1993. "Family Structure and the Risk of a Premarital Birth." *American Sociological Review* 58(2):210-32.
Zajonc, Robert B. 1976. "Family Configuration and Intelligence." *Science* 192:227-36.
Zajonc, Robert B., and Gregory B. Markus. 1975. "Birth order and Intellectual Development." *Psychological Review* 82:74-88.