

DISSERTATION

SOCIAL DETERMINANTS OF COLLEGE COMPLETION AND WEALTH MOBILITY: A  
LIFE COURSE APPROACH TO EDUCATIONAL COMPLETION AMONG YOUNG BABY  
BOOMERS

Submitted by

Matt Aronson

Department of Sociology

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Fall 2015

Doctoral Committee:

Advisor: Michael Lacy

Jennifer Cross  
Paul Stretesky  
David Most

Copyright by Matt Aronson 2015

All Rights Reserved

## ABSTRACT

### SOCIAL DETERMINANTS OF COLLEGE COMPLETION AND WEALTH MOBILITY: A LIFE COURSE APPROACH TO EDUCATIONAL COMPLETION AMONG YOUNG BABY BOOMERS

This dissertation fits within both the sociology of education subfield and the mainstream discipline's longstanding concern over stratification, which is generally understood as the systematic persistence of unequal social statuses and socioeconomic positions in society. Educational status, especially the condition of having completed a four-year undergraduate degree, represents a key feature of socioeconomic position and an important predictor of various life chances in the United States. One limitation of previous sociological research on education is that most studies have asked about the *whether* of educational attainment without giving much attention to the *when* of degree completion. A main goal of this dissertation, then, is to remedy that inattention by asking how family-level and individual-level conditions during people's childhoods may influence the timing of their four-year college completions. Another goal is to examine wealth mobility, and to ask whether timely college completion influences wealth mobility from early- to mid-adulthood. I offer three essays, each of which addresses these goals in different ways. The first study compares results from an event history model of high school completion and argues for treating educational completion as an event in time rather than as a binary or categorical outcome. That section offers a methodological contribution to current scholarship as sociologists of education are increasingly taking advantage of longitudinal data sets. The next essay asks how some characteristics and conditions early in individuals' lives may influence the timing of their college completions. I consider teenage childbearing, family

poverty, maternal education and other factors in an attempt to provide an alternative to the conventional understanding of race-ethnicity as a predictor of individuals' conditional odds of college completion. The fourth essay departs from the emphasis on "timing of college completion" as an outcome and instead focuses on several under-examined questions about wealth mobility (movement within the wealth distribution over the life course) and whether timely college completion and adolescent employment are associated with upward wealth mobility over the adult life course. The final section also makes a basic contribution to social scientists' understanding of wealth dynamics among the National Longitudinal Survey of Youth's late baby boomer cohorts.

## ACKNOWLEDGEMENTS

I extend special thanks to my mentor and good friend, the venerable Mike Lacy. Among many other things, he taught me the value of asking good sociological questions. Much gratitude as well to my friend and pedagogical guru, Jeni Cross, for encouraging the teacher inside me. I am eternally grateful to my family, and especially to my wonderful Paul, for everything.

## TABLE OF CONTENTS

ABSTRACT.....	ii
ACKNOWLEDGEMENTS .....	iv
SECTION 1: INTRODUCTION .....	1
Time and Life Course Theory .....	2
Trends and Policy Contexts in U.S. Education.....	6
The Dissertation’s Organization .....	8
The Data.....	8
Methodological Details and Caveats .....	9
General Note About Survival Analysis.....	12
The Basic Arguments.....	15
SECTION 2: TREATING EDUCATIONAL COMPLETION AS A TEMPORAL EVENT .....	17
Data, Measures, and Methods .....	21
Sample Statistics and Binary Response Models .....	26
An Event History Approach to High School Completion .....	30
Predicted Probabilities .....	36
Discussion and Conclusion .....	41
SECTION 3: SOCIAL ORIGINS, EARLY PARENTHOOD, AND RISK OF COLLEGE COMPLETION AMONG YOUNG BABY BOOMERS .....	45
A Life Course Perspective on Education .....	47
Kinds of Time .....	48
Research on Stratification and College Completion.....	50
Data, Methods, and Measures .....	54
Research Questions .....	61
Sample Statistics and Regression Models.....	62
Regression Models for Conditional Odds of College Completion .....	73
Concluding Discussion and Ideas for Future Work.....	99
SECTION 4: “BETTER TO DO (SOME) THINGS EARLY”: TIMELY 4-YEAR COLLEGE COMPLETION, TEEN PARENTHOOD, ADOLESCENT EMPLOYMENT, AND WEALTH MOBILITY IN MID-ADULTHOOD.....	107
Focus on Wealth .....	110
Research Questions .....	116
Data and Methods .....	118
Variables .....	118
Sample Statistics and Descriptive Analyses .....	123
Logit Regression Models of Wealth Persistence and Mobility .....	133
Concluding Discussion and Future Work .....	139
REFERENCES .....	144
APPENDIX A.....	156
APPENDIX B .....	160
APPENDIX C .....	163
APPENDIX D.....	168

## SECTION 1

### INTRODUCTION

This dissertation fits within both the sociology of education subfield and the mainstream discipline's longstanding concern over stratification, which is generally understood as the systematic persistence of unequal social statuses and socioeconomic positions in society. Social scientists have long studied formal education in industrialized societies, particularly during recent decades in the United States as high school and college have become ever more necessary for upward social and economic mobility. Educational status, especially the condition of having completed a four-year undergraduate degree, represents a key feature of socioeconomic position and an important predictor of various life chances. One limitation of previous sociological research on education is that most studies have asked about the *whether* of educational attainment—as either an outcome or an independent variable—without giving much attention to the *when* of degree completion. A main goal of this dissertation, then, is to remedy that inattention by asking how family- and individual-level conditions during people's childhoods may influence the timing of their four-year college completions. Another goal is to examine wealth mobility, and to ask whether timely college completion may influence wealth mobility during mid-adulthood. To these ends I offer three essays, each of which addresses the goals in different ways. The first study (Section 2) compares results from an event history model of high school completion, which examines the timing of high school degree completion, to a conventional logistic regression approach with a binary outcome. That essay argues for treating educational completion as an event in time rather than as a binary or categorical outcome, and offers a methodological contribution to current scholarship as sociologists of education are increasingly taking advantage of longitudinal data sets. The next essay (Section 3) adds to the

scholarship on educational stratification by asking how some individual- and family-level characteristics and conditions early in individuals' lives affects the timing of their college completions. I consider teenage childbearing, family poverty, maternal education and other factors in an attempt to provide an alternative to the conventional understanding of race-ethnicity as a predictor of individuals' conditional odds of college completion. The fourth essay (Section 4) departs from the emphasis on "timing of college completion" as an outcome and instead focuses on several under-examined questions about wealth mobility (movement within the wealth distribution over the life course) and whether timely college completion and adolescent employment are associated with upward wealth mobility over the adult life course. Section 4 also makes a basic contribution to social scientists' understanding of wealth dynamics among the National Longitudinal Survey of Youth's late baby boomer cohorts.

### **Time and Life Course Theory**

Generally speaking, my work emphasizes some key tenets in the sociology of education and the mainstream discipline's concern with social stratification. Sociologists of education study many things, from education's links to health and occupational prestige to the meanings of educational status and its association with processes of endogamy, social mobility, friendship homophily, and so forth. Conventional wisdom holds that education, whether it is conceptualized as predictor or outcome, is crucial to so many dimensions of people's life chances (e.g., Pallas 1993; Bowles and Gintis 2003). The matter of at what ages people complete high school or a four-year college degree has become the focus of some recent research, but it is still an underexplored issue. That the timing of educational completion has only recently come to scholars' attention is somewhat surprising given the venerable tradition of research from the "life



course” perspective, which emphasizes age and era as crucial factors in examinations of role entries and exits, status transitions, and other time-varying phenomena. With this to mind, in each of the manuscript’s sections I attempt to modestly enhance our understanding of timely educational completion, its social determinants, and its association with certain aspects of people’s socioeconomic fates such as adult wealth mobility. I analyze each of these issues with life course theory as a common thread across the sections.<sup>1</sup>

The methodological issues discussed below connect substantively to life course theory and its general focus on timing across lives rather than snapshot views of expected transition ages or periods. Specifically this dissertation’s sections are rooted implicitly (and sometimes explicitly) in the “timing of lives” concept from life course theory, which posits three kinds of time—*individual*, *generational*, and *historical* time (Price, McKenry, and Murphy 2000; also see Elder 1985 and Edler 1998 for classic statements on the life course perspective). Individual or biographical time refers to what may seem intuitive, namely the cultural significance of chronological age and age periods, a key assumption being that the timing of conditions, events, and transitions in people’s earlier lives can influence their social positions in later life. So-called “timely” transitions such as completion of high school around age 18 or of four-year college around age 22 are social constructions based on culturally shared age definitions and understandings of normative social timetables, and they become patterned according to the social meaning attached to chronological age markers (Hagestad and Neugarten 1985).

Generational time refers to age cohort, as in the discrete generations we name such as “baby boom generation,” “generation X,” “Millenials,” and so forth. Cultural norms and

---

<sup>1</sup> The terms “life course perspective,” “life course approach,” and “life course theory” will be used interchangeably throughout because I am convinced that there is such a thing as life course *theory* as much as there is a life course approach or perspective (e.g., Elder et al. 2003). I made this decision while also being aware of the debate among sociological theorists as to whether the life course perspective constitutes a cohesive theory or merely an orienting perspective.

meanings associated with certain life stages, or the ages at which people are expected to move from one condition to another, certainly differ across cohorts. For example, high school completion is now very common, a modal status transition in late adolescence, but only for relatively recent American generations. Today it is perhaps mundane to say that high school completion (diploma or GED) is a profoundly meaningful goal to most people, a socially shared expectation necessary for respectability or for the expected status passages into adulthood and working life. By contrast, high school completion was much less normative and less meaningful as a key life event or status marker for members of generations who grew up in the late 1800s through the mid-1900s, eras in which most people experienced status distinction as a function of their family heritage and personal work biographies rather than their formal educational achievements.

Relatedly, historical time refers to societal changes, major economic shifts, or other generally notable events and how these affect people's lives. In terms of theorizing a phenomenon like educational completion among late baby boomers, a historical time perspective would emphasize the effects of World War II and the post-war economy because of their widespread and lasting societal effects. Perhaps also of interest would be the unprecedented mid-century growth of higher education institutions and federal education policy, major shifts in industrial labor markets, and other historical changes that created the conditions for high school and college completion to become more normative and more common in people's lives.

Since individuals' past and current situations are influenced by historical and social contexts, often their personal origins and outcomes are strongly influenced by the ascriptive power of social inheritance; that is, just as a daughter may eventually inherit her parents' wealth, a son may eventually "inherit" social things such his single mother's low educational status or

inability to move upward along the wealth ladder during later adulthood. Also of interest is how individuals' statuses (a person's wealth status, for example) can persist or change across their working lives. In short, a long tradition of research has demonstrated that advantaged and disadvantaged statuses can persist across and within generations (e.g., Erikson and Goldthorpe 1992; Shavit and Blossfeld 1993; Conley 2001). Part of the current work attempts to theorize and examine empirically people's early-life conditions in their families of origin and certain ascribed characteristics as predictive of age at college completion and wealth mobility in adulthood. As explored in Section 3 below, of particular interest are the varied routes through which, say, racial minority status (or socioeconomic disadvantage in one's family of origin, or mother's education, etc.), may affect individuals' educational outcomes and their timing. Later in Section 4, I will analyze within-person or intra-generational wealth mobility, and attempt to discern how this phenomenon varies by the ages at which individuals' complete college, whether they engaged in paid work during adolescence, whether they became parents during their teenage years, and so forth.

The various kinds of time—individual, generational, and historical—may be thought of as distinct yet related, even overlapping in certain respects. For each section in this dissertation, individuals and individual time will serve as the primary unit and level of analysis, respectively, although there will be some mention of historical time with the intra-generational wealth analyses in Section 4. However, in the brief summarizing discussion of the general trends and policy contexts in United States education presented below, I attempt to account for all three kinds of time.

## **Trends and Policy Contexts in U.S. Education**

A key premise in life course theory is that people generally follow a series of age-normative transitions (specified ages or age periods at which people are expected to begin voting, working, drinking, attending high school, engaging in sexual activity, etc.), and that such age-graded patterns will change according to shifts in a society's cultural norms, institutional arrangements, economic conditions, and so forth. In other words, people's subjective experiences in individual time are embedded in historical and generational time. While in the current work I do not formally examine macro-level trends or their links to the notions of historical and generational time, it is important to acknowledge some of the historical context for college completion and its antecedents in K-12 schooling.

Florencia Torche (2011, 767), in her examination of cohort trends for age-adjusted college completion rates among people born between 1905 and 1965, identified a "period expansion" in college completion for those born between the 1910s and the late 1940s, a "period decline" in completion for those born in the 1950s, and another expansion for late Baby Boomers born in the 1960s. Reasons for the expansions and declines in college completion include growth in earnings returns to schooling, variation in federal investments toward higher education, legalized racial segregation, the emergence of the G.I. Bill, and sudden availability of college draft deferments. It was among the generations that came of age during the first two-thirds of the twentieth century that college completion became established as "age-normative" around the 20- to 22-year old mark. People born during the late Baby Boom (ending in 1965) were born into the tail end of this period, thus one question is how much heterogeneity in age-at-completion we should expect among them. Considering that the rise of a substantial non-traditional aged student population did not begin until late in the twentieth century, one would expect relatively little

heterogeneity among late Baby Boomers, i.e., that they would have typical ages at completion similar to those of earlier generations.

Considered at another level, formal education in the United States reflects various tensions between the cultural ideals of equal opportunity, merit-based achievement, and socioeconomic mobility, all of which are embedded in the various types of time discussed previously. Structural phenomena related to group mobility and achievement at the population level evoke the generational and historical notions of time. For instance, it is plausible that if in the 1970s the U.S underwent a major shift toward, say, widespread public support for funding early childhood education programs or college readiness curricula in public high schools, then we could have expected stronger labor markets and greater national economic stability today and into the future (Heckman 2001). Or if, in earlier times, school-level racial segregation had been less persistent since the beginning of publicly funded education, perhaps we could expect there to be narrower racial disparities in college completion today. I mention these historical counterfactuals to suggest that past events and processes matter for the dynamics of college completion today, and suffice to say that this relatively simple premise informs the current study.

A four-year college degree can help individuals overcome the constraints of disadvantaged origins (it can loosen the ascriptive bonds of family background, so to speak), but college also plays a role in altering the level of socioeconomic inequality at the societal level, especially given the degree to which labor markets reward college attainment relative to how they reward high school diplomas. As will be discussed later on, four-year college is getting less affordable, more normative (but still class- and race-differentiated), and more valuable. There are voluminous scholarly literatures on many of the issues just discussed; these issues comprise a

general context for the relevance of my own work, in addition to the specific literatures referenced in each of the forthcoming sections.

### **The Dissertation's Organization**

Instead of the traditional chapter-based progression, this manuscript is divided into four distinct parts, with sections 2, 3, and 4 written in the style of a conventional journal article (though each section may be lengthier than some journals allow). Excepting the current introduction, the three substantive sections are meant to stand as “studies on their own” more or less, with a life course theoretical orientation serving as the common thread. These sections are Section 2, a modeling comparison (methodological) study; Section 3) an analysis of how the conditional odds of individuals’ four-year college completions may be influenced by family background, parental education, ascriptive features of people lives, sibling characteristics, adolescent childbearing, and early familial income; and finally Section 4, which details the dynamics of intra-generational wealth mobility and volatility, and whether earlier college completion may boost people’s chances for upward wealth mobility. While each section has a distinctive substantive and analytical focus, all sections share some common features which I discuss below.

### **Data**

In order to best capture people’s early life conditions, family background, educational fates, and movement up and down the wealth distribution over time, I took advantage of a well-known panel dataset that provides detailed information and that followed its respondents over many years of their lives. Panel data have several advantages over cross-sectional survey data, and these advantages will be mentioned in the discussions to come. With this in mind, each of the forthcoming sections uses data from a nationally representative sample of Americans who came

of age during the late 1970s and early 1980s. The samples were derived from the National Longitudinal Survey of Youth 1979 cohort (NLSY), a complex panel survey that originally included a nationally representative sample of 12,686 men and women aged 14 to 21 years on December 31, 1978. The NLSY is a multistage, stratified, clustered probability sample meant to represent the entire population of youth residing in the United States in 1979. The sample respondents were re-interviewed annually through 1994 and biennially thereafter, and I used data for the time period 1979 to 2004. The NLSY has important advantages for an event history analysis because data about the timing of educational attainment and many of its potential predictors were collected contemporaneously at each survey wave, which gives rich detail and enhances accuracy. In each section below, the samples were restricted to respondents who remained in the survey over a period of 25 years, from 1979 to 2004. (For more detailed discussion of the NLSY and descriptions of all variables used in the subsequent analyses, see *Appendix A* at the end of this document. )

## **Methodological Details and Caveats**

### *The Metric for Time, and a Note on the Interpretation of Statistical Models*

In each section of this dissertation, the primary units of analysis are individuals who were between the ages of 14 and 22 as of December 31, 1979. Ideally, the most sensible metric for time should reflect the cadence that one expects to be most useful for the outcomes of interest. Since I am interested in tracking outcomes at certain points in respondents' life courses, "respondent's age" in a given year will represent time. Researchers who study educational transitions usually measure time discretely because the events under study only occur, or are only assessed, at regular intervals every week, month, semester or year (Willett and Singer

1991). For reasons of conceptual clarity, I assume “year” (or age) as the discrete case in each section in this dissertation.

Throughout each section, the words “associated with,” “correlated,” or “predictive of” are used so as not to imply that the data represent “proof” of causal relationships between independent and dependent variables (even though the data relationships may in fact be causal in social reality). And although the discussions below emphasize substantive significance regarding the relationships between variables, statistical significance is reported for coefficients indicated in the models.<sup>2</sup>

Besides the decomposition models that test for mediation effects as presented in Section 3, all coefficients will be presented as odds ratios, which represent exponentiated slope coefficients (in logistic regression, odds ratios are the exponential of the log-odds ratio). As has been well discussed in the methodological literature (Menard 2002, 57), the odds ratio is one among several ways to address the generic question “How does a one-unit change in an independent variable ( $X$ ) multiplicatively change the predicted odds of some event ( $Y$ )?” Since this dissertation employs an event history framework, “the odds of some event ( $Y$ )” is not the same as “the odds of some event ( $Y$ ) ever occurring.” Instead, and usually implied yet sometimes explicitly stated, the “odds of an event” will refer to the instantaneous odds (discrete-time odds) that individuals experienced an event at age  $t$  given that they had not yet experienced the event up to that point.

---

<sup>2</sup> Statistical significance was evaluated in the usual manner, namely with test statistics to generate probabilities (p-values) that observed relationships in the samples would appear as they do if there truly (in social reality) were no relationships between predictor and outcome variables. That being said, I did not rely solely on p-values in interpreting the associations between variables.



### *Measurement Issues and Overcoming the “Window Problem”*

The analyses in this dissertation rely on several decades of panel survey data gathered by the United States Bureau of Labor Statistics. Survey data provide a set of rigorously collected, well-tested, and meaningful markers that must be used in place of perfect information about the individuals surveyed. Of substantive significance for the current study, panel data like NLSY offer a distinct advantage when it comes to dealing with the so-called “window problem” in survey research.

Many researchers who study educational completion with retrospective data have been able to measure events or circumstances for only a short duration or window during respondents’ childhoods. So-called “window variables” are commonly accepted as proxies for information that is not otherwise available but that presumably characterizes someone’s entire childhood or at least a substantial time period. (The indicator “lived with both biological parents at age 14” is a commonly used window variable in sociology; for various reasons, researchers have had to assume that it approximates the existence of an “intact” or two-parent family across a respondent’s entire childhood.) However, variables that describe intermittent events or discontinuous periods of more persistent characteristics may fail to closely correspond to an actual childhood experience. As methodologists have shown, window variables often are poor proxies because they may be inconsistent with information measured over a person’s entire childhood, adolescence, or adulthood (Wolfe et al. 1996). The NLSY, a panel data set, has substantial advantages including that it offers a more temporally comprehensive view of many conditions across respondents lives, not just during a single year or limited range of years. For instance, I was able to determine specifically when respondents had completed college instead of whether they “completed college by age 25” or some such window variable.

### *Language Used in Decomposition Models of Direct, Indirect, and Full Effects*

Regarding the decomposition models presented in Section 3, it is appropriate to include some comments about the program and language used therein. The Stata add-on *khb* (Karlson et al. 2012; Breen et al. 2013) was used to detect mediation effects. The decomposition process requires breaking down a given model's total effects into the sum of direct and indirect effects.<sup>3</sup> In principle, assessing mediation requires the consideration of three key relationships: 1) The observed association between a given predictor and the outcome. 2) The relationship between a given predictor and the mediator variable. 3) The association between the mediator variable and the outcome. The *khb* program gives direct, and indirect effects in nonlinear regression models as reported below, and was helpful in assessing, for example, whether racial differences in discrete-time odds of college completion are attributable to the effect of teenage parenthood. As seen in the models presented later in Section 3, the term “Diff” refers to the calculated difference between the predictor's total and direct effects on the outcome (it represents the indirect effect of being Black as it operates through the mechanism of teenage parenthood, with discrete-time college completion as the outcome), the term “Full” indicates the direct effect including control variables, and “Reduced” refers to the total effect without control variables.

### **General Note About Survival Analysis**

Survival analysis is about analyzing how certain predictors affect the time-to-occurrence of an event. This method has been used widely in studies of important life events such as occupational changes, onset of sexual activity, age at first marriage, childbirth, and so forth. “Event

---

<sup>3</sup> This is but a curt reference to recent methodological advances, the computational details of which are beyond the scope of the current work, which have enabled the accurate estimation and interpretation of total, direct, and indirect effects in nonlinear regression models (see Karlson et al. 2012 and Breen et al. 2013).

occurrence” refers to a person’s transition from one condition to another condition.<sup>4</sup> Survival models are especially useful because they allow researchers to treat “time-to-event” as a variable, whatever the event or condition may be. For example, using educational completion as a dependent variable, it is sensible to define college completion as the condition or “event” of interest—a respondent has either experienced the event of college graduation at time  $t$  or they have not yet completed. The measure for “time” is central in studies concerning whether and when events occur in people’s lives (Singer and Willett 2007), and the measure should reflect a cadence that one expects to be most useful for the outcome of interest. I chose respondent’s age in years to represent time, which simply means I measured time discretely in yearlong intervals. The concept of event-time refers to when in respondents’ lives such events occur, and as mentioned above this section examines how certain predictors affect the timing of respondents’ college completion.

Survival analysis is useful because it helps in dealing with problems that arise when the data violate normality assumptions that usually underlie linear regression. When analyzing the “time-to-occurrence” of an event—for example, analyzing whether respondents graduate from college before a certain age cutoff—the distribution for time until college completion may be non-normal. In situations where the distribution of time until an event is non-normal (nonsymmetric, bi-modal, etc), traditional linear regression is unable to deal with this violation of normality (Cleves et al. 2008, 3). Survival analysis handles this problem because it does not require the usual normality assumptions about time-to-event distributions.

---

<sup>4</sup> As it is currently understood, survival analysis uses various types of estimation to trace cases over time, enabling researchers to estimate respondents’ likelihood of transitioning from one “state” or condition to another. Most broadly, this method derives from methodological developments in disciplines ranging from the biomedical sciences to epidemiology, sociology, and educational research.

One component of an event history analysis is the survivor function, which is a cumulative statistic that indicates the probability that an individual will “survive” in the dataset (i.e., not experience a given event) beyond a certain time. In ordinary language, this refers to the probability that a respondent will survive in the risk set beyond time  $t$ , that is, that she will *not* experience the event of interest beyond a certain time  $t$ , thus remaining or “surviving” in the dataset. The survivor function is of interest here because it estimates the fraction of the population that has yet to complete a given level of education in each successive time period, and its value indicates the proportion of people exposed to each period’s hazard (Singer and Willett 2003, 334). At the point in time when no respondents have yet experienced college completion (say age 16 to be on the safe side), then everyone is said to be “surviving” in the data set and the survivor function’s value is exactly 1. Over time, however, as people complete college the survivor function declines toward zero. Once someone experiences an event, or is dropped from the risk set, then the statistics program no longer considers that individual in its computations.<sup>5</sup> Any respondent’s survival probability for college completion is the probability that the individual will not have completed college beyond time period  $t$ . For this to be the case and in order to speak of “conditional” probabilities, a given individual must not have experienced the event during time period  $t$  or during any earlier period. (Survival analysis has some counterintuitive terms: What researchers often technically call a “failure at time  $t$ ” may in real life represent something desirable such as someone’s college completion, but it can be called failure because the respondent “did not survive in the risk set” beyond time  $t$ .)

Another feature is the *hazard function*, which gives a statistical summary of event occurrence and is a key underlying dependent variable in survival analysis. When using discrete

---

<sup>5</sup> The survivor function can only stay the same or decrease, it will never increase. For example, when passing through time periods when no respondents complete college, the survivor function stays at its previous level.

time data, the hazard rate is the probability that a person will experience an event at time  $t$  while the person is still “at risk” of experiencing that event. Unlike the survivor function, which cumulates information across time, the hazard function expresses the conditional probability or unique risk of experiencing an event associated with each time period, so it is an unobserved variable that reflects both the occurrence and the timing of events (Singer and Willett 2003, 330). This can reveal the time period(s) when events are especially likely or unlikely to occur.<sup>6</sup>

### **The Basic Arguments**

Here, I give a brief overview of the particular questions addressed in each section of the dissertation. First, the discussion and modeling comparisons presented in Section 2 attempt to address an issue in the sociology of education scholarship, namely that conventional point-in-time measurements of education such as “years of schooling” (or, as is often done categorically, “highest degree attained” or “completed degree by age 25”) overlooks the heterogeneity in the ages at which individuals complete their schooling. Further, as a substantive matter, in most societies the timing of people’s status transitions is crucial socially and economically, so I argue that studying educational completion as an event in time can help researchers to better model how and when educational completion may matter for people’s socioeconomic fates over the life course. This particular comparison compares results from an event history model of high school completion, which examines the timing of high school degree completion, to a conventional logistic regression approach with a binary outcome for educational completion. In short, the

---

<sup>6</sup> To illustrate with college graduation as the event of interest: Imagine a fourteen-year old girl’s hazard rate is 1.2 for college graduation at time  $t$ , and her 30-year old cousin has a hazard of 2.2 at time  $t$ . The younger girl has half the risk of having ever completed college (given that she had not completed by time  $t$ ) compared to the older girl’s risk. Intuitively we would expect most college graduations in the US to occur during respondents’ early adult years.

results support treating educational completion as an event in time rather than as a binary outcome.

Next, Section 3 addresses some questions relevant for the educational stratification scholarship. First, I describe the distribution of and sample statistics for individuals' ages at four-year college completion. Then, I specify survival models of college completion using life course theory and recent stratification scholarship as a guide. Third, based on decomposition analyses of potential mediation effects, I argue for the importance of considering teenage childbearing, family poverty, maternal education, and other variables as they help to explain the apparent negative effect of minority race-ethnicity on odds of college completion; this provides a subtle alternative to the conventional understanding of race-ethnicity as a predictor of college completion.

Finally, in Section 4 I attempt to extend our understanding of wealth dynamics among individuals in the NLSY by first interpreting some descriptive results and addressing questions of intra-generational wealth mobility from early adulthood to mid-adulthood, and then by examining racial differences in wealth mobility over that same period. The basic but important finding is that some things (adolescent employment and timely college completion) but not others (teenage parenthood) seem to portend better chances for people to experience upward wealth mobility.

## SECTION 2

### **TREATING EDUCATIONAL COMPLETION AS A TEMPORAL EVENT: FORMAL COMPARISON OF EVENT HISTORY AND BINARY RESPONSE MODELS<sup>7</sup>**

I argue for treating educational completion as an event in time rather than as a binary or categorical outcome. This section offers a methodological contribution to current scholarship as sociologists of education have increasingly taken advantage of longitudinal data sets (as distinct from multiyear cross-sectional data sets) to address empirical questions. Using panel data from the NLSY, I compare results from an event history model of high school completion, which examines the timing of high school degree completion, to a conventional logistic regression approach with a binary outcome. Although results from the conventional approach and event history analysis are similar in certain respects, the latter has several advantages. As a substantive matter, the timing of educational completion affects other later life outcomes, so adopting a method that recognizes the timing of educational attainment is preferable in principle. Second, my findings show that the event history approach yielded slightly weaker coefficients but narrower confidence intervals for predictors across most models. Third, comparing the predicted probabilities of ever completing high school shows that the event history perspective more closely approximates actual event occurrence. Finally, event histories can better incorporate time-varying predictors than can binary outcome models, allowing researchers more sophisticated ways to model the effects of various life experiences.

Like so many socioeconomic and cultural phenomena, educational completion has long interested researchers as an outcome to be explained and as a predictor of other outcomes.

---

<sup>7</sup> An earlier version of material in this section was presented as a paper at the Midwest Sociological Society's annual meetings (Aronson and Lacy 2014).

Recently critics have decried social scientists' "failure to replicate" in various substantive and methodological domains, and my contribution here, while not replicating any particular study's findings, was meant to formalize what sociologists are increasingly assuming as axiomatic about how to properly model educational completion. The conventional approach has been to treat educational completion dichotomously, and although education researchers have paid some attention to the timing of educational completions, in sociology there has been too little emphasis on the ages by which people achieve educational milestones. Dichotomization in logistic modeling can obscure researchers' knowledge about educational transitions because it eliminates potentially meaningful variation in event times (Willett and Singer 1991, 408). That is, dichotomization does not distinguish respondents who completed high school at age 18 from those who completed at age 21 or 27, etc. It is possible to address the question "What factors predict the timely completion of high school?" with traditional logistic modeling by simply constructing a categorical age-graded measure for completion that discerns whether respondents completed high school before their twentieth birthday, before their twenty-fifth birthday, and so forth, and as noted this is fairly commonly used by sociologists of education (Haas and Fosse 2008).

An even more detailed potential solution using traditional regression would be to categorize "age at completion" so that the possible outcome values include many discrete ages (17, 18, 19, 20, etc.). Constructing variables in either of these two ways may seem like a sensible way to trace the timing of people's educational careers. However, for statistical reasons neither approach is optimal because standard regression approaches cannot systematically address the dilemma of "censoring" implicit in panel data, namely the question of what to do with



respondents who were not observed to have experienced a given event during the data collection period (Singer and Willett 2003, 316).

Inspired by life course scholarship that emphasizes trajectories and transitions over time (e.g., Pallas 1993; Furstenburg 2003; Wu 2003; Pallas 2004), this section formally explores whether there is analytical benefit to examining educational completion in truly longitudinal fashion. Specifically, I compare an event history approach and a binary response model for analyzing educational attainment in the National Longitudinal Survey of Youth 1979 (NLSY).

As noted, models of a non-timed educational outcome can be useful, but degree timing deserves more attention than because it matters greatly for socioeconomic fates in the U.S., with consequence beyond the simple fact of completion or non-completion. Graduating with a high school diploma at age 18 has different social and socioeconomic consequences than completing a GED at age 33. Conversely, the social origins of people who completed college later than, say, age 25, may differ systematically compared to those who completed at more “normative” ages. Some status attainment research has formalized these prosaic observations; completing high school at “non-traditional” ages tends to reduce people’s chances of enrolling and completing college (see for example, Mare 1980; Shavit and Blossfield 1993), and research has shown that timely academic attainment predicts various socioeconomic opportunities later in life (see Frisco 2008 for a summary). If the timing of education and related status transitions had remained relatively homogeneous in the United States, off-time sequences might affect very few people and thus not matter much; however, today fewer young adults adhere to the “traditional” sequence of finishing school, getting a job, marrying and having children (Mouw 2005). Greater variation in education timing could prefigure greater variation in socioeconomic outcomes. While the issue of timing has interested social scientists who study life course transitions such as

family formation, career entry/exit, and the like (Frase 1989; Jacobs and King 2002; Maralani 2011), the increasing heterogeneity of educational trajectories in the United States means that timing deserves more attention in studies of high school and college careers.

Much research has usefully considered the “when” of event occurrence in various domains (Upchurch and McCarthy 1990; Conger et al. 1994; Murnane et al. 1999), yet there has been surprisingly little sociological research in regard to the timing of educational completion (*cf.* Jacobs and Berkowitz-King 2002) despite what many researchers may intuitively sense about it. While the broader motivation of the entire dissertation is a substantive interest in how educational timing is affected by factors such as parental wealth and poverty during childhood, the current section examines a way to incorporate timing and so better model educational completion.

I use the National Longitudinal Survey of Youth 1979 cohort to compare two modeling approaches for high school completion: one treating it as a dichotomous outcome, and the other treating it as an event with an indicator of age-at-occurrence. I compare results from a binary logistic regression to those from a discrete-time event history method with proportional odds. While other researchers have noted the general desirability of using event history methods to analyze educational attainment (e.g., Singer and Willett 2007; Jacobs and Berkowitz-King 2002), I know of no previous work that directly compares these two approaches on the same sample. Both approaches include some basic background factors, including race-ethnicity/ethnicity, sex, teenage childbearing, parental education, and family poverty as predictors of high school completion. Compared to a conventional binary approach, the event history method revealed a similar though subtly richer picture of this phenomenon.

## **Data, Measures, and Methods**

The study sample comprises respondents from 26 waves of the National Longitudinal Survey of Youth 1979 (hereafter NLSY), a nationally representative sample of 12,686 persons aged 14 to 21 when first enrolled in the survey beginning in 1979. The NLSY is a multistage, stratified, clustered probability sample meant to represent the entire population of youth residing in the United States in 1979. The sample respondents were re-interviewed annually through 1994 and biennially thereafter, and I used data for the time period 1979 to 2004. The NLSY has important advantages for an event history analysis because data about the timing of educational attainment and many of its potential predictors were collected contemporaneously at each survey wave, giving rich detail as well as enhancing accuracy. During the survey's most recent years, over 9,000 individuals were still participating in the survey.

For the analyses presented here, NLSY's special military sample was excluded because the majority of these respondents were not interviewed after 1983. Also excluded were non-interviews and observations with missing data on the independent variables. This has been done in other studies using NLSY data, e.g., see Klepinger et al. 1995; Garasky 1995.<sup>8</sup> Due to these constraints, the final sample for the current analyses comprised 6,155 respondents who were followed over the same 25 year period, 1979 to 2004. Educational completion rates of people in the sample were consistent with high school completion rates in the United States generally; among the sample respondents, 86% ( $n = 5,288$ ) completed high school at some point during the follow-up period, with 14% ( $n = 867$ ) never completing.

---

<sup>8</sup> I did not use sampling weights here. Comparisons of various models with and without weights, personal consultation with Dr. Jay Zagorsky (Zagorsky 2010), and my reading of NLSY technical documentation (NLSY 1983) led to this decision.

Two separate analyses of the data are performed here. One is a conventional logistic regression approach, for which the unit of analysis was the individual respondent and the outcome of interest was whether s/he ever completed high school. In this straightforward approach using a person-oriented dataset, each person has one record (or case) of data. The other approach is a survival approach using a person-period data structure in which each person has many records depending on “how long they lasted in the risk set.” In both modeling situations, I defined the risk period for high school completion as running from age 14 to 47, as this corresponded to the lowest and highest values in the sample.

### *Outcome Measures*

High school completion is here defined as the reported attainment of a high school degree or GED equivalent at any point prior to age 47. (The sample was truncated in this way because so few individuals completed high school beyond their mid-40s.) The outcome variable was whether a respondent completed high school or GED equivalent during the observation period (the “risk period” that occurred between 14 and 47 years of age).<sup>9</sup> This corresponds to common past practice in which researchers document differences between people who achieve a certain educational attainment and those who do not, irrespective of when in a person’s lifetime such attainment occurs (e.g., Mare 1981; Grubb 1993; Kane and Rouse 1995; Chen and Kaplan 1999). In the event history portion below, the age at which this happened was also noted.

---

<sup>9</sup> It would have been possible to measure degree attainment by distinguishing respondents who completed high school at, say, age 18 from those who completed at age 21, while still using a binary response model. However, since the goal was comparison it seemed sensible to use a relatively simple binary measure (this was done for other methodological reasons as well, described below).

### *Explanatory Measures*

The explanatory variables were essentially identical for the event history and logistic regression models.<sup>10</sup> I included gender, given the finding that women are more likely than men to complete high school (Goldschmidt and Wang 1999). I included indicators of racial/ethnic status for black and Hispanic, with white as the reference category.<sup>11</sup> Previous research has found black and Hispanic Americans to have lower high school graduation rates, though this likely is due to a combination of family-based differences in cultural capital endowment as well as socioeconomic and family background factors that happen to be associated with race-ethnicity (Hauser, Simmons, and Pager 2004). Indeed, studies that adjust for the differences in cultural and economic endowments that blacks and whites “bring with them” into the educational system have shown the black high school completion rate to be comparable to that of non-Hispanic whites (Conley 1999). Regardless, I included racial/ethnic status considering the general importance of racial/ethnic status in U.S. society.

Parental financial resources influence children’s academic performance and educational attainment (McLanahan and Sandefur 1994), and they may reflect important differences in the kinds of cultural capital children inherit from their parents (Lareau 2003). Financial resources obviously shape other opportunities such as the ability to provide enriching after-school educational experiences that influence academic achievement and later economic well-being in adulthood (Duncan 1994). Researchers commonly examine the effects of family socioeconomic resources, so I included an indicator for whether respondents lived below the federal poverty line during at least one of their childhood years while living in their family’s household.

---

<sup>10</sup> NLSY79 respondents were first interviewed at various ages ranging from 14 to 22, so the binary response models controlled for number of years respondents could reasonably have been eligible to complete high school, beginning with age 14.

<sup>11</sup> As have most other studies that use NLSY79, I omitted “Asian” and “American Indian” racial classifications due to substantial limitations in the initial sample.

Additionally, based on conventional wisdom in social mobility research, I expect parental educational attainment to strongly predict children's chances of ever completing high school (Garasky 1995). Thus, I included two parental education variables for mothers and fathers, respectively, to indicate their total years of formal schooling. In my sample, this ranges from 4 to 21 years.

Ever since Moynihan's famous report on the potential "pathologies" of certain family structures, researchers, policy makers, and the lay public have noted the various disadvantages associated with childbirths to teenage parents, especially to unattached teenage women (Moore and Waite 1977; Furstenburg et al. 1989; Furstenburg 2003), so I included "early childbearing" as an explanatory measure. I recognize that the presumed causal connection between teenage parenthood and parental disadvantage is vigorously debated among social scientists (see for example, Geronimus et al. 1999; Luker 1996; Lawler et al. 2001; Geronimus 2003). Some researchers argue that, net of various background and personal characteristics, having a child while in high school does not significantly increase a mother's risk of dropout, and other studies have suggested that unmeasured family background characteristics may attenuate the presumed disadvantages of teen childbearing (Hoffman, Foster, and Furstenberg 1993; Hoffman 1998).<sup>12</sup> In sum, researchers have not definitively discerned the causal influence of teen childbearing per se on a person's chances of high school completion, but addressing this controversy is beyond the scope of my study. Given the recognized importance of this variable, I chose to include it, with teenage parenthood defined as ever becoming a biological parent child before age 20.

---

<sup>12</sup> Personal characteristics other than teenage fertility (but that may be closely associated with it) such as age at sexual initiation, parental characteristics, adolescent "problem behaviors," and so forth have been shown to negatively affect high school completion, controlling for teen parenthood's effect (Upchurch and McCarthy 1990; Francesconi 2007; Frisco 2008).

## Methods

Here I offer some brief comments on event history models, which as the name implies were designed to make time-until-event the focus of analysis. Event history or survival models with longitudinal data offer a more sophisticated treatment that accounts for the timing of various phenomena (Mayer and Tuma 1990; Singer and Willett 2003).<sup>13</sup> Although the substantive advantages of this analytic approach are of primary interest here, it also has important statistical advantages regarding issues such as error distributions and censoring, as is well-described in those standard sources.

To specify, my event history approach models the log-odds of high school completion (event occurrence), using the following form:

$$\ln\left(\frac{P_{it}}{1-P_{it}}\right) = \alpha + f(t) + \sum_{k=1}^K \beta_k X_{ikt}$$

$P_{it}$  is the hazard, or the probability that the  $i^{th}$  respondent completes high school at time  $t$  given that s/he has yet completed.  $\alpha$  is a constant,  $f(t)$  is a function of time,  $X_{ikt}$  is the value of the  $k^{th}$  covariate at time  $t$  for respondent  $i$ , and the  $\beta_k$  are coefficients associated with the  $k$  different explanatory variables.<sup>14</sup> Unless explicitly modeled differently, the assumption is that a variable's effect (e.g., gender's effect) is the same at all points during the risk period, so I omitted  $t$  as a subscript on the  $\beta$ s.

The most direct interpretation of the estimates for such a model is in terms of the exponentiated  $\beta$  coefficients, which measure the multiplicative effect of the explanatory variables on the odds of high school completion at any given point in time (Box-Steffensmeier

---

<sup>13</sup> The event history approach has gained widespread popularity in mainstream sociology, and has been used in studies of life events such as occupational mobility, onset of sexual activity, family formation, children's educational achievement, educational progress, and so forth (e.g., Jacobs and King 2002; Bahr 2009; Wagmiller et al. 2010; Roksa and Velez 2012).

<sup>14</sup> Here I estimated parameters with a conventional approach with the NLSY data in person-period format using a binary logistic regression model. (e.g., see Singer and Willett 2003; Allison 1992.)

and Jones 2004). I also use this model equation to generate hazard functions, showing the predicted hazard at any point in time, for a particular combination of predictor values. Additionally, I examine survival functions, which indicate the predicted probability of “surviving” in the data set (not having yet completed high school) and how it might vary over time.

### **Sample Statistics and Binary Response Models**

Here I present descriptive statistics for the sample and basic results from the standard logistic regression models for high school completion without treating it as an event in time.

Subsequently I will compare these results with those from logistic regression models using an event history approach.

Table 2.1 displays descriptive statistics for all respondents for whom there was valid information on the explanatory variables during a 25-year observation period. As expected, most individuals completed high school at some point, with almost three-quarters completing prior to age 19, and with the proportion of females completing by that age exceeding that for males. The proportion of White respondents completing prior to age 19 was more than double that of Hispanic and Black respondents. These statistics are in line with Census and Current Population Survey data that indicate similar sex and racial disparities as well as upward trending rates of high school completion in the late twentieth century (Chapman et al. 2011). Lower completion rates were observed for respondents whose families experienced at least one year of poverty and for those who became parents as teenagers. About two-thirds of respondents whose mothers had



at least 12 years of formal education completed high school prior to age 19, which is 26 percentage points greater than respondents whose mothers had completed less than 12 years.

**Table 2.1: Descriptive Statistics**

<b>Variables</b>	<b>% Completed High School</b>	<b>% Completed H.S. Prior to Age 19</b>
<b>Sex</b>		
Male (n = 3,115)	83.6	47.6
Female (n = 3,040)	88.2	58.6
<b>Race-ethnicity</b>		
White (n = 3,731)	86.8	59.0
Black (n = 1,404)	88.4	22.2
Hispanic (n = 1,020)	79.2	22.9
<b>Teen Parent</b>		
Yes (n = 1,054)	71.8	34.6
No (n = 5,101)	88.8	56.9
<b>Family Poverty</b>		
Yes (n = 2,416)	76.7	39.8
No (n = 3,739)	91.9	61.6
<b>Mother Yrs. of Education</b>		
≥ 12 yrs. (n = 3,492)	93.1	64.2
< 12 yrs. (n = 2,663)	76.5	38.1
N = 6,155	85.9	72.7

Table 2.2 presents results from multiple models, beginning with only race-ethnicity and gender as the first set of independent variables, and adding others in subsequent models. This table shows how various explanatory variables (teenage parenthood, poverty experience, mother's educational attainment) add to or help interpret race-ethnicity and gender as predictors of degree completion. Starting with the initial model in the first column of Table 2.2, my results showed that females had about 1.5 times the odds of ever completing high school as males, Hispanic respondents about half the odds of whites, and Black respondents differed little from whites (note the considerably wider confidence intervals for black respondents, however). Being female was associated with higher odds of completing high school than males across all four

model specifications, with the relative odds rising as various factors (e.g., teen parent status) were brought into the model.

**Table 2.2**  
**Odds Ratios for High School Completion Using Logistic Regression**

Variables	Model 1	Model 2	Model 3	Model 4
<b>Black</b>	0.943 (0.776 - 1.147)	1.135 (0.927 - 1.389)	1.452*** (1.177 - 1.793)	1.594*** (1.285 - 1.978)
<b>Hispanic</b>	0.461*** (0.382 - 0.555)	0.509*** (0.420 - 0.617)	0.618*** (0.507 - 0.754)	1.268* (1.000 - 1.608)
<b>Female</b>	1.457*** (1.257 - 1.689)	1.892*** (1.616 - 2.216)	1.900*** (1.618 - 2.230)	1.961*** (1.663 - 2.313)
<b>Teen parent</b>		0.267*** (0.224 - 0.317)	0.318*** (0.266 - 0.380)	0.336*** (0.280 - 0.403)
<b>Fam. Poverty</b>			0.367*** (0.312 - 0.431)	0.517*** (0.436 - 0.614)
<b>Mother's edu.</b>				1.101*** (1.066 - 1.137)
<b>Father's edu.</b>				1.117*** (1.089 - 1.146)
Constant	1.243 (0.951 - 1.624)	1.459** (1.109 - 1.920)	3.129*** (2.288 - 4.277)	0.307*** (0.196 - 0.482)
observations	6,155	6,155	6,155	6,155
Chi-square	198.3	408.8	560.6	792.5
Df	4	5	6	8
BIC	-48853	-49055	-49198	-49412
Craig-Uhler R2	0.0570	0.115	0.156	0.217

confidence intervals in parentheses

\*\*\* p<0.001, \* p<0.05

The coefficients and associated confidence intervals in Model 1 indicate that being black was associated with only a negligible decrease in respondents' chances of completing high school, but in all three of the subsequent models blacks' odds of completion were higher compared to whites. This may be due to the inclusion of a relatively crude measure for family poverty as an indicator of respondents' class origins; neighborhood socioeconomic status and family economic resources (parental wealth, especially) have been shown to account for much of the black-white

and Hispanic-white disparities in educational completion and other important outcomes (Conley 1999; Lareau and Conley 2008).<sup>15</sup>

The task here is model comparison rather than formal empirical analysis, so a more prosaic result is the positive association between parental education and the odds of degree completion. In Model 4 of Table 2.2, higher parental educational attainment is associated with greater odds of respondents ever having completed high school in their lifetimes—in the sample, an additional year of father’s education is associated with an increase in the odds of completing high school by a factor of 1.2. Granted, this association may reflect whatever *direct* effect parental education has on high school completion in addition to its correlation with other variables such as socioeconomic status, parenting style, parental expectations, household environment, individual aspirations, peer influences, and many other characteristics not included here. Conventional social science wisdom would indicate that parental education is substantially and positively associated with the odds of respondents’ educational completions.

Additionally, having a child prior to age 20 was associated with a decrease in the odds of ever completing high school by a factor of 0.34, adjusting for other variables, as compared to respondents who had no biological children before that age.<sup>16</sup> Research indicates that this association may be confounded by other family background factors and adolescent behaviors (not accounted for here) common among individuals who are most likely to become teenage

---

<sup>15</sup> Another possible explanation for these coefficients, based on the practice of “social promotion” in K-12 education (National Research Council 1999; Hauser 1999), is that black students may be concentrated in schools where teachers have tended to promote racial minorities into higher grade levels due either to internal administrative pressures or out of sympathy for the challenges such students face in daily life instead of promoting based on academic performance or readiness (see Hauser 1999 and Lorence et al. 2002). Further, it is notable that research on social promotion is very sparse for time periods predating the emergence of social promotion as a major public concern in the 1990s, thus we cannot know whether social promotion was more or less prevalent during the time when these individuals were in their early school years.

<sup>16</sup> As has become well recognized recently (see Williams 2009 and Karlson et al. 2012), directly comparing and testing coefficients across groups (gender, race-ethnicity, etc.) in logit models requires special care, particularly when interpreting interaction effects. Since the purpose here is *model* comparison, I did not examine interaction terms in these models.

parents. These factors include age at sexual initiation, parents' aspirations, adolescent "problem behaviors," and so forth (Upchurch and McCarthy 1990; Hoffman, Foster, and Furstenberg 1993; Hoffman 1998; Francesconi 2007; Frisco 2008). But the goal here is not to elaborate on the explanatory power of teenage parenthood, parental education, race-ethnicity, or gender vis a vis respondents' high school completions. Rather, as shown below, what I see as notable is how this model's results compare to those yielded in the second modeling approach.

### **An Event History Approach to High School Completion**

Having inspected some basic results from the standard logistic regression model with high school completion as a binary outcome, I used the same sample to examine models derived from logistic regression for event history data. With high school completion treated as an event in time, the following tables and figures present some descriptive statistics involving the survivor function, cumulative hazard, and hazard function, each of which will be explained below. The ensuing discussion then argues that event history modeling reveals a slightly different view of high school completion compared to the models discussed earlier. Finally, I will explore other characteristics of the event history approach.

When considering possibilities for approximating the outcome variable in my event history models, it was necessary to regress high school completion on age. Initially the most general specification had every discrete age as a predictor (14, 15, 16, and so forth, up to age 47), but it was more parsimonious to collapse the very few early completers into a single category.<sup>17</sup> The "risk period" for this sample began at age 15, the youngest age at which more than one

---

<sup>17</sup> Among these possibilities, I compared goodness-of-fit, values for the Bayesian Information Criterion, etc.—using many discrete ages at first, and then using various age-range categories—in order to discern the optimal modeling form to represent time-to-completion.

respondent completed high school, and ended with the completion of high school or age 46 (or the point at which a given individual was “censored” in the data set).

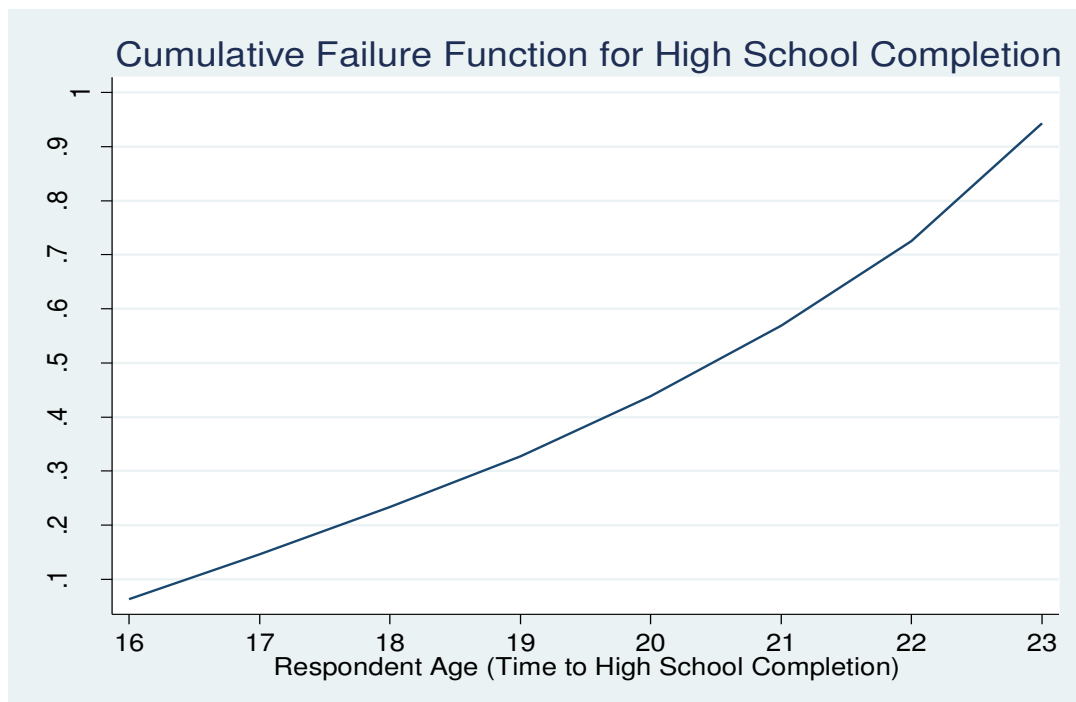
Table 2.3 displays the survivor function and cumulative incidence of completion for the sample. The survivor function cumulates the period-by-period proportion of respondents who remain in the risk set, i.e., the proportion of respondents who had not yet completed high school by each age period. As might be expected, the proportion of respondents remaining “at risk” of completing high school drops precipitously beyond age 22. (By definition the survivor function can never increase, so the proportion of respondents who remain “at risk” of completing high school continues to decline incrementally into the later ages beyond those shown in the table.<sup>18</sup>) The cumulative incidence, also known as cumulative failure, simply cumulates the proportion of respondents who were estimated to have ever completed high school by the given age intervals. There was only a very small amount of “remaining risk” by the time these individuals reached their early twenties; Table 2.3 shows that, among respondents who ever completed high school during their lifetimes, approximately 94 percent of them were estimated to have done so before age 23.

**Table 2.3**  
**Survival and Cumulative Incidence for High School Completion**

<u>Age Interval</u>	<u>Survival</u>	<u>Cumulative Incidence</u>
15 – 16	0.937	0.063
16 – 17	0.854	0.146
17 – 18	0.766	0.234
18 – 19	0.673	0.373
19 – 20	0.562	0.438
20 – 21	0.431	0.569
21 – 22	0.275	0.725
22 – 23	0.057	0.943

<sup>18</sup> Due to censoring, and because some respondents never completed high school during the entire data collection period, the estimated survivor function does not fall to zero.

Since high school completion is a normatively age-graded life transition (normative in the sense that, when people do complete high school, almost everyone completes prior to their early twenties), then it is easy to conceptualize it as something that becomes less and less likely as people get older. This is a prosaic way of saying that, in a survival framework, the likelihood of dropping out of the risk set (i.e., completing one's schooling) cumulates over time, to a certain point or age after which completion becomes very unlikely. To visualize this idea, Figure 2.1 displays a graph of information from the righthand column in Table 2.3, or the cumulative incidence (or “cumulative failure”) of high school completion. Like Table 2.3 but in graphical form, Figure 2.1 shows that nearly all respondents who ever completed high school during their lifetimes were estimated to have done so by age 23, beyond which there is very little “remaining risk” of individuals completing high school.

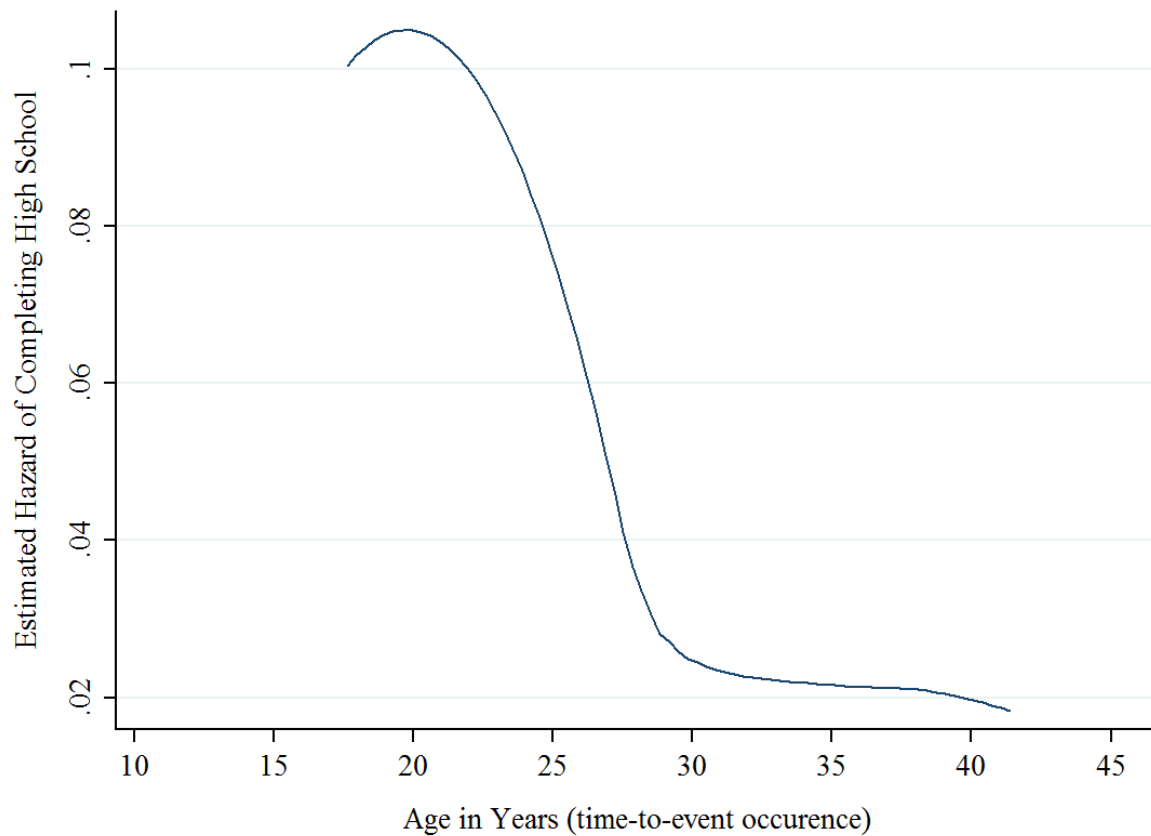


**Figure 2.1**  
**Incidence of High School Completion by Age**

A more sensitive tool for describing the “whether and when” of high school completion is to graph the set of estimated hazard probabilities at each age. Unlike the survivor function, which cumulates information across time, hazard assesses the unique and conditional risk of event occurrence at each age (Singer and Willett 2003, 341). Figure 2.2 presents a graph of the values for discrete-time hazard of high school completion associated with each age. Here, the hazard function represents the conditional probability of completing high school in each age period among respondents who had not completed in prior age periods; the greater the hazard in a given time period, the greater the estimated “risk” that eligible respondents would complete high school during that time period. As presented in Figure 2.2, the highest chances of completing high school occur prior to age 20, with 17 to 20 being the period of greatest risk. As might be expected, the hazard function peaks during respondents’ late teens and sharply declines thereafter.<sup>19</sup>

---

<sup>19</sup> Unlike the survivor function, hazard estimates can increase or decrease in value over time; the fact that hazard does not increase beyond age 20 reflects the rigidly age-graded nature of high school completion in the U.S. population.



**Figure 2.2**  
**Discrete-Time Hazard Estimate for High School Completion**

Table 2.4 displays odds ratios for the same set of respondents as presented in Table 2.2, but it differs from the previous regression models because now the outcome variable reflects age at high school completion for all respondents. As in the discussions above, Table 2.4 presents several models, except here they indicate how the independent variables successively add to race-ethnicity and gender as predictors of the *timing* of high school completion rather than simply the odds of ever completing. Compared to the standard logistic regression models that used a binary outcome, the models presented in this survival approach reflect individuals' conditional "discrete-time odds" of completing high school at any given age.



The binary response and event history modeling scenarios differ sufficiently as to make direct coefficient comparisons inappropriate, so instead I focus on the most relevant points of comparison, namely within-model changes (or persistence) in direction and magnitude of coefficients. Starting with the model presented in column one, the event history results showed that women had about 1.3 times the odds of high school completion at any age compared to men, a relationship that increases slightly across all four models in Table 2.4 as more covariates are added.<sup>20</sup> This result is identical in trend and direction to the result for gender in the previous standard logistic regressions using a binary outcome. However, as displayed in Table 2.4, there is a weaker association between individuals' conditional odds of high school completion at given ages and all predictors (e.g., looking at model 4 in Table 2.4 and model 4 in Table 2.2 shown previously, the odds-ratio for female here is 1.589 compared to 1.961), and the event history models yielded narrower confidence intervals. It is important to note that the coefficients reflect different things in each modeling scenario: In the binary response models, they represent an effect or association cumulated across the entire risk period, which we would expect to be much larger than the effect at any particular point in time (as is measured in the event history model).

**Table 2.4**  
**Odds Ratios for High School Completion Using Event History Analysis**

Variables	Model 1	Model 2	Model 3	Model 4
<b>Black</b>	0.946 (0.865 - 1.034)	1.041 (0.950 - 1.140)	1.193*** (1.086 - 1.309)	1.263*** (1.149 - 1.388)
<b>Hispanic</b>	0.599*** (0.542 - 0.661)	0.622*** (0.563 - 0.687)	0.681*** (0.615 - 0.753)	1.028 (0.920 - 1.148)
<b>Female</b>	1.290*** (1.198 - 1.388)	1.563*** (1.447 - 1.688)	1.563*** (1.446 - 1.689)	1.589*** (1.469 - 1.720)
<b>Teen parent</b>		0.406*** (0.367 - 0.448)	0.460*** (0.416 - 0.509)	0.479*** (0.432 - 0.530)
<b>Fam. Poverty</b>			0.531***	0.665***

<sup>20</sup> The caveat discussed in footnote 9 above (regarding the inappropriateness of comparing coefficients across models in logistic regression) also applies here.

			(0.490 - 0.574)	(0.613 - 0.722)
<b>Mother's edu.</b>				1.081***
				(1.063 - 1.099)
<b>Father's edu.</b>				1.072***
				(1.058 - 1.086)
Constant	1.153*** (1.076 - 1.236)	1.191*** (1.110 - 1.278)	1.431*** (1.328 - 1.541)	0.240*** (0.198 - 0.292)
Observations	47,890	47,890	47,890	47,890
Chi-square	13941	14280	14532	14952
Df	7	8	9	11
BIC	-497142	-497470	-497711	-498109
Craig-Uhler R2	0.509	0.520	0.528	0.541

Confidence intervals in parentheses  
\*\*\* p<0.001

## Predicted Probabilities

As mentioned previously, these two modeling approaches predict different aspects of the outcome, and for various reasons there is no straightforward way to directly compare what each one predicts. So as to attempt to directly comparing the results from the two approaches, I used each to generate “predicted probabilities” of ever completing high school. I obtained the predicted probabilities of high school completion from the logistic regression model shown in Table 2.2, but for survival modeling a different technique is needed; the event history model equation was used to predict the probability of event non-occurrence at each age across the entire risk period. Specifically, following Singer and Willet (2003), the predicted probability that an individual did not graduate from high school by the end of her last period of observation  $T$  is

$\sum_{t=1}^T (1 - \hat{P}_{it})$ , where  $\hat{P}_{it}$  is the model-predicted hazard of high school completion for respondent  $i$  at time  $t$ . This was then used to calculate the predicted probability that an individual *did* complete high school during the period of observation using  $1 - \sum_{t=1}^T (1 - \hat{P}_{it})$ . The two models, by different

mechanisms, can arrive at a prediction for the same outcome, and seeing whether they differ provides another way to assess the relative usefulness of each approach.

To do this, I first computed predicted probabilities for the full sample from the standard logistic model (model 4 in Table 2.2) and those from the survival analysis model (model 4 in Table 2.4). Scoring actual high school completion as 1 or 0, I then computed the difference for each approach between the predicted proportions of occurrence and actual proportions of respondents completing high school. Table 5 displays a basic view of how the two approaches differ, and presents mean algebraic and absolute differences between actual event occurrence and predicted probabilities of completion obtained from the event history and standard logistic regression approaches.

Additionally, since it is more informative to examine predictions at specific values (Long and Freese 2006)—it is conventional to use means, medians, or other purposely specified values—Table 2.5 displays the differences in predicted probability of completing high school for several “archetypal” cases, or individuals (selected for the purpose of comparison) who had specified values on some of the independent variables. The archetypes here compared two racial identifications (*white* and *black*) for women, as well as whether those respondents had a biological child during their teenage years (*Teen Parent*), whether they experienced at least one year of poverty in their families of origin (*Family Poverty*), and whether respondents’ mothers had less than 12 years of education (*Mother’s Education*).

As indicated in Table 2.5, the event history and standard logistic regression approaches yielded very similar predicted probabilities for the full sample. Overall, predictions from the binary outcome model were only slightly closer to the observed (actual) proportion of respondents who completed high school. The “mean algebraic and absolute difference from

observed” calculations simply represent algebraic and absolute mean differences, respectively, between the proportion of respondents who were predicted to have completed high school and the proportion of respondents who actually completed at some point in their lives. Mean algebraic differences indicate the general direction of error in the predictions for high school completion, and mean absolute differences summarize the overall accuracy of the predictions. Compared to the predictions based on the logistic model with a binary outcome, the predictions obtained from an event history approach were slightly more accurate for three of the four archetypes (archetypes 2, 3, and 4); the “mean absolute differences from observed” were closer to the observed or actual proportion of respondents completing high school. While these comparisons do not *definitively* proclaim one or the other approach’s superiority, they do suggest that event history-based predictions may more closely approximate actual event occurrence for high school completion.

**Table 2.5**  
**Comparing Observed High School Completion with Predicted Probabilities of Completion from Binary Logistic and Event History Models, Using Four Respondent Types<sup>a</sup>**

Group	Observed Proportion Completing H.S.	Predicted Proportion Completing H.S. (Mean Algebraic Difference from Observed) [Mean Absolute Difference from Observed]*	
		Event History Model	Binary Logistic Model
Full Sample	0.859	0.838 (-0.021) [0.006]	0.842 (-0.017) [0.004]
Archetype 1	0.976	0.951 (-0.025) [0.013]	0.957 (-0.019) [0.011]
Archetype 2	0.489	0.468 (-0.021) [0.019]	0.527 (0.038) [0.023]
Archetype 3	0.958	0.912 (-0.046) [0.041]	0.901 (-0.057) [0.049]
Archetype 4	0.638	0.556 (-0.082) [0.142]	0.522 (-0.116) [0.162]

<sup>a</sup>The four respondent types were specified as follows:

Archetype 1 : white female: never teen parent, no family poverty, mother’s education greater or equal to than 12 yrs  
Archetype 2 : white female: teen parent, family lived in poverty at least a year, mother’s education less than 12 yrs  
Archetype 3: black female: never teen parent, no family poverty, mother’s education greater than or equal to 12 yrs  
Archetype 4 : black female: teen parent, family lived in poverty, mother’s education less than 12 yrs

\* For both modeling scenarios, the “mean algebraic and absolute difference from observed” calculations simply represent algebraic and absolute mean differences, respectively, between the proportion of respondents predicted to have completed high school and the proportion of respondents who actually completed.

Another way to discern differences between the two modeling approaches is to examine and compare the association or effect of each explanatory variable between them. Because the coefficients measure different things, they needed to be standardized to make them comparable. I did this re-expressing the odds ratio for each explanatory variable as a multiple of another odds ratio in the model. For the two modeling situations shown in Table 2.6, this was done using two different odds ratios as a base for comparison, first with the odds ratio for *female* as the base and then again using that for *mother's education*. This table shows that the effects of most predictor variables, relative to the odds ratio for *female* and the odds ratio for *mother's education*, were more modest (closer to 1) in the event history framework. For example, in the logistic model, the effect of *teen parent* on high school completion relative to the effect of *female* was 1.52, but only 1.31 in the event history model. There were some exceptions to this; for example, the effect of *mother's education*, as a proportion of the effect of *female*, was slightly stronger in the event history model compared to the standard logistic model. I did not detect radical differences in the relative effects of variables across the two approaches, but they do differ enough such that the modeling choice is not trivial.

**Table 2.6**  
**Comparing the Effects of Primary Predictors on the Odds of High School Completion**  
**(Controlling for Gender and Mother's Education)**

Logistic Regression vs. Event History: Relative Effects of Primary Predictors, *Female* as Reference

Variable	Raw Odds Ratios		... as a Proportion of <i>Female</i> Effect	
	Logistic	Event History	Logistic	Event History
<b>Female</b>	1.96	1.59	1.00	1.00
<b>Teen Parent</b>	0.34	0.48	1.52	1.31
<b>Family Poverty</b>	0.52	0.67	0.99	0.95
<b>Mother's Edu</b>	1.10	1.08	0.56	0.68
<b>Black</b>	1.59	1.26	0.81	0.79
<b>Hispanic</b>	1.27	1.03	0.65	0.65

Logistic Regression vs. Event History: Relative Effects of Primary Predictors, *Mother's Education* as Reference

Variable	Raw Odds Ratios		...as a Proportion of <i>Mother's Education</i> Effect	
	Logistic	Event History	Logistic	Event History
<b>Female</b>	1.96	1.59	1.78	1.47
<b>Teen Parent</b>	0.34	0.48	2.70	1.93
<b>Family Poverty</b>	0.52	0.67	1.76	1.39
<b>Mother's Edu</b>	1.10	1.08	1.00	1.00
<b>Black</b>	1.59	1.26	1.45	1.17
<b>Hispanic</b>	1.27	1.03	1.15	0.95

## **Discussion and Conclusion**

Modeling educational completion as a dichotomous outcome using a binary response model is simpler than using event history, and the coefficients derived from both approaches were similar in direction and relative magnitude. So why bother with the complexity of an event history model? My answer may resonate for many researchers interested in the timing of educational completion (or the timing of school enrollment, dropout, or transfer) as outcomes and as predictors of other later-life outcomes. Conventional logistic regression and event history analysis are both capable of modeling high school completion and its predictors in the basic sense that I demonstrated; the conventional logistic and event history approaches yielded fit statistics with similar shifts in magnitude across all four model specifications (the respective R-squared statistics steadily increased and corresponding BIC statistics declined substantially over the successive models). The overall story is that the event history results appear similar compared to results derived from standard logistic regression with a binary outcome, which is unsurprising considering my relatively simple model specifications. However, I noticed three differences between the approaches: First, compared to the modeling approach using a binary outcome, the event history approach yielded coefficients that were slightly weaker for all predictors and across most of the regression models. Second, the event history approach gave more precise estimates, with narrower confidence intervals for every independent measure and across all model specifications. Third, when I estimated respondents' predicted chances of ever completing high school, the event history and logistic frameworks yielded less similar results, with the event history predictions offering a slightly more accurate approximation of actual event occurrence for high school completion.

As indicated in the introductory comments, my results support what many researchers probably sense logically or feel intuitively, namely the value of studying various aspects of the life course as temporal processes. This matters because event history analysis allows for more sophisticated outcome variables, enabling researchers to address more nuanced research questions that take fuller account of the *whether and when* of event occurrence instead of merely the *whether* (Wu 2003). Simply completing a degree may matter less than the age at which people do so, as has been generally indicated by researchers studying the differences between conventional high school completion and GED attainment (e.g., Heckman and Rubinstein 2001). Beyond the so-called diploma-GED divide, there may be more nuanced differences between those who earn a diploma at, say, age 17 compared to those who earn a diploma at age 18 or 19. One plausible substantive example is that timely educational completion may interact with the timing and durations of people's experiences in the labor market; methodologically, this may resonate with researchers who study occupational mobility, work trajectories, and socioeconomic mobility. Indeed, the often indirect effects of "off-time" educational completion are ripe for further study. Recently, researchers have shown that later high school completion portends later college enrollment, and "off-time" college enrollment undesirably affects eventual college completion net of various pre-college background characteristics (Bozick and DeLuca 2005; Goldrick-Rab and Han 2011). In short, the results here lend methodological support for what seems to be the intuitive value of approaching educational completion as an event in time.

Examining even relatively simple descriptive results using an event history framework, as with Table 2.3 and Figure 2.1 above, revealed useful information such as the time periods when events are especially likely or especially unlikely to occur, i.e., the periods of greatest "risk" at which respondents undergo life course transitions. It can also allow researchers to model other



covariates' effects on time-varying outcomes such as educational completion or childbearing. Another possibility is that some covariates that are themselves time-varying could also have time-varying effects on a time-varying outcome like school completion. For example, one hypothetical scenario is that respondents who lived in low-income households during their teenage years may be delayed in high school, whereas respondents who lived in low-income households during only their first five years of life might be more "on time," or vice versa.

Some recent work has treated educational completion as an event in time, but more is yet to be done. For example, Wao (2010) illustrated the general benefit of multilevel discrete-time analysis in his study of graduate degree completion, and Roksa and Velez (2012) recently demonstrated the usefulness of discrete-time hazard modeling in their study of the timing of college completion. There are many other interesting modeling possibilities that take account of the timing of family formation, labor force participation, parenthood, educational re-entry, etc., and indeed substantial efforts have been made in these regards (see Elman and O'Rand 2004; House et al. 2005; Wao 2010; and Roksa and Velez 2012). For sociologists of education, the model comparisons discussed here point up the value of using survival analysis to study educational attainment as a temporal process in the life course rather than merely a step on people's "status attainment" ladders. Since quantitatively-oriented sociologists typically have not adequately examined people's educational careers as temporal processes, my analysis supports the axiom from life course research that there is substantial benefit in doing so (Elman and O'Rand 2004; O'Rand et al. 2009). Education may be cause or consequence of important life course trajectories, transitions, conditions, and spells; and with regard to its influence on socioeconomic trajectories and durations, approaches of this kind are becoming more possible as the increasing availability of rich longitudinal survey datasets (NLSY, Panel Study of Income

Dynamics, the National Education Longitudinal Survey, the National Survey of Households and Families, and others) permits truly longitudinal pictures of people's lives.

The basic methodological point here has simple but very practical implications; event history modeling allows a more informative view of the timing of educational completions, and the longitudinal treatment of education as "completions in time" enables researchers to examine the effects of other factors associated with the timing of completion in more detail compared to conventional or person-level logistic approaches. While I do not claim that event history is the *only* useful approach to analyzing educational degree outcomes, I do suggest it is optimal for researchers wanting to acknowledge the fundamentally longitudinal character of educational completion.

## SECTION 3

### SOCIAL ORIGINS, EARLY PARENTHOOD, AND RISK OF COLLEGE COMPLETION AMONG YOUNG BABYBOOMERS

#### Introduction

How do family conditions during adolescence, ascriptive factors such as race-ethnicity, gender, or siblings' education, and individual behaviors such as teenage childbearing affect the ages at which people complete college? Sociologists have carefully documented the apparent connections among family background, gender, race-ethnicity, and education, and they have examined changes within these relationships across cohorts and within cohorts over time (Warren et al., 2002, 432; DiPrete and Grusky 1990; Grusky 2001). But due to what for a long time was an inadequate availability of longitudinal data and limitations in statistical modeling techniques, much research in the human capital, social origins, ascriptive stratification, and status attainment traditions focused on the *level* of education rather than the timing and temporal patterning of education as important factors in stratification processes. This focus on level of education (and not the timing of education) is understandable given the current context in which those who are most successful in labor markets—those with job security, nonstagnant wages, retirement portfolios, and so forth—tend to have post-secondary levels of education compared to those who are less successful. Since the mid-2000s however, researchers have developed a better appreciation for timing in degree attainment and its potential importance for status stratification over the adult life course. Some of this research has emphasized the influence of individuals' childhood family structures and financial resources on timely college completion, and other work has focused on the later-life consequences of timely college completion, especially the

connections between timeliness and higher wages (see Elman and O’Rand 2004 for more detailed critique of these limitations).

The work presented in this section contributes to sociological scholarship on educational stratification by asking how some individual- and family-level characteristics and conditions early in individuals’ lives affect the odds of their completing college at any given age. Elman and O’Rand (2004) have argued that the strong connection between timely college completion and adult socioeconomic outcomes (wages, primarily) is explained mostly by differences in parental resources, childhood family structure, teenage parenthood, and ascribed characteristics such as race-ethnicity. While I share an interest in these factors, the fundamental issue in this section is what their influence may be on educational timing *per se* instead of its well-established association with later-life outcomes such as adulthood wages. Specifically, using a nationally representative sample of late baby boomers born between 1957 and 1965 (NLSY 1979) and an event history modeling framework, I attempt to do three things. First, I describe the distribution of and sample statistics for individuals’ ages at four-year college completion. Then, I specify survival models of time-to-completion using life course theory and recent stratification scholarship as a guide. Third, with a decomposition analysis of potential mediation effects, I argue for the importance of considering teenage childbearing, family poverty, maternal education, and other variables; this provides an alternative to the conventional understanding of race-ethnicity and gender as predictors of individuals’ conditional odds (“discrete-time” odds) of college completion.

## **A Life Course Perspective on Education**

Since the 1970s post-secondary education has become a key part of ever more Americans' lives, and social scientists have noted the socioeconomic importance of completing at least some college or a four-year degree. By the year 1950, just over a third (34 percent) of Americans had completed high school, but by the year 2000 more than 80 percent had done so. Of particular interest for the current work is that the share of American adults who had completed *college* quintupled from just over 5 percent in 1950 to 25 percent as of 2010 (U.S. Census Bureau 2010). College completion is becoming more normative, and there is little doubt that higher levels of education positively affect socioeconomic outcomes (Pallas 2003, 166). Decades of social science research has demonstrated that not finishing high school or college is associated with lower lifetime earnings, lower occupational attainment, lower family income, and greater likelihood of welfare receipt and material deprivation during adulthood (Mueller and Cooper 1986; Veum and Weiss 1993; Garasky 1995; Duncan et al. 2004). Adults with more years of formal education are more likely to participate in the labor force during their lifetimes compared to those with less education (Bound et al. 1995). As well, education reliably predicts occupational status (Featherman and Hauser 1978; Grusky and DiPrete 1990), more schooling tends to translate into higher lifetime earnings, and heads of household with college degrees have greater net worth compared to those with lower levels of education (Sewell and Hauser 1980; Murnane et al. 1995; Land and Russell 1996).

The current work addresses issues related to educational timing rather than educational attainment per se and is motivated by life course theory, which emphasizes timing (of events, transitions, conditions, etc.) across lives rather than snapshot views of expected transition ages. A guiding premise in life course theory is that people's lives generally follow a series of age-

graded or age-normative transitions, such as fertility or educational completion. These often have patterned trajectories (Rauscher 2011, 554; Sampson and Laub 1992) that change according to historical trends and shifts in a society's cultural norms, institutional arrangements, economic conditions, and so forth. Timely transitions such as high school completion around age 18 or college completion around age 22 are based on normative timetables that become patterned according to the social meanings of chronological age (Hagestad and Neugarten 1985). These patterns are more than analytical constructs that social scientists created to aid categorization and analysis; research has shown that people are meaningfully oriented toward various "social clocks" that correspond with societal shifts (Furstenburg 2003; O'Rand 1996, 2002), hence the assumption that there is such a thing as being ahead of or behind what is conventionally seen as a normative transition age (or age range) for a given life event such as education, and that there are real consequences—of both the socioeconomic and cultural kind—for being ahead or behind.

### **Kinds of Time**

Among the three kinds of time posited in life course theory—*individual*, *generational*, and *historical*—the first two are most relevant here. Individual time refers to the cultural significance of people's numeric age and the age periods that are deemed meaningful. A basic assumption is that "timely" transitions such as completion of four-year college around age 22 are social constructions based on culturally conventional age definitions and shared understandings of normative social timetables (Hagestad and Neugarten 1985; Price, McKenry, and Murphy 2000). A more complex assumption, one that life course researchers have used to address empirical questions, is that the timing of conditions, events, and transitions in people's earlier lives may influence their social positions in later life, and that the timing of one transition may condition

the timing or likelihood of others. For example, teenage childbearing may, through various pathways, reduce a person's chances of completing college before their mid-20s, which in turn may reduce one's lifetime earning potential or reduce one's upward wealth mobility later during the peak earning years. As well, a person's age at first childbirth directly and indirectly alters one's likelihood of the timely achievement of other transitions; teenage childbearing often leads to school drop-out, and it can boost a person's chances of being "ahead of time" on other transitions such as adolescent employment or marriage which, in turn, may further dampen people's chances of educational completion.<sup>21</sup> (These are but a few of the many logical routes people often take when arguing about "precocious childbearing.") Some previous research confirms these intuitions. For example, Pallas (1993) found that people who are "off-time" (older than age 23) in completing a four-year undergraduate degree tend to come from less socioeconomically advantaged families, and they tend to be less socioeconomically successful in adulthood compared to those who completed college earlier. In a study arguing that "the race-ethnicity is to the swift," Elman and O'Rand (2004) showed that earlier college completion portends higher lifetime earnings.

Being off-time in educational completion can also be somewhat stigmatizing. Earning a bachelor's degree at age 29 may be just as emotionally rewarding and worthy of celebration as earning one at age 22, but within the normative context of socioeconomic achievement in U.S. society, the later completion is associated with "off-time" status and its attendant disadvantages such as less time to earn at higher levels, diminished cultural capital, and so forth. With an

---

<sup>21</sup> As has been widely examined in the sociological literature, a principle that characterizes stratification within cohorts over time is the so-called "Matthew effect," a process whereby achievements earlier in life enable people to more rapidly acquire valued resources and social positions that yield cumulative returns over time. While I do not test for a Matthew effect in terms of cumulative socioeconomic returns to timely educational completion (in the form of labor market earnings or occupational status trajectories), this general principle does resonate with the tenets of life course theory that motivate my research.

outcome like educational completion, a general expectation from life course theory is that college degrees earned at later ages will tend to reflect familial socioeconomic disadvantage and many other factors such as race-ethnicity, gender, family structure, the proximal influences of significant others such as siblings, and individual behaviors such as adolescent childbearing or college aspirations.

The second kind of time, generational time, refers to the distinct generations we name such as “Baby Boomers,” “Millennials,” and so forth. Cultural norms and meanings associated with certain life stages, or the ages at which people are expected to move from one condition to another, certainly differ across generations. Compared to their forebears, Baby Boomers grew up in a time of substantial federal support for higher education, public educational expansion, and relative economic prosperity. And even today, four-year college completion is much more common than it was just half a century ago. Between 1965 and 2005, the share of American men ages 30-60 who have a bachelor’s degree more than doubled from 9 percent to 19 percent, and among same-aged women it went up even further, from 7 percent to nearly 20 percent (Torche 2011). Norms have shifted such that college is now seen as necessary for substantial upward mobility and, in the eyes of many, perhaps even for passage into respectable middle class adulthood.

### **Research on Stratification and College Completion**

Relatively little previous scholarship has focused on timing of college completion, but considerable research has examined the broad role of college completion in society. A four-year college degree has become both a meritocratic “equalizer” and a significant stratifier in the United States. It is an equalizer in terms of its role in enabling socioeconomic achievements in



adulthood that are not determined by family origins. The U.S. is unique among nations in that the direct impact of parental resources on children's eventual economic attainments is extremely weak (virtually zero "intergenerational transmission" of socioeconomic background) among children who complete college compared to those who ended their educations with high school (Hout 1988; Torche 2011). This is one of the most striking findings in stratification research, and it suggests that for those who earn a four-year degree, their socioeconomic position in adulthood is likely to be independent of their family's socioeconomic origins.

A four-year college degree may serve a crucial meritocratic role in enabling people to overcome the ascriptive power of disadvantaged origins, but it also represents a stratifying mechanism in terms of societal-level inequalities. Earning a college degree matters greatly in industrialized "credential societies" characterized by universal K-12 where adults have easier access to post-secondary education compared to their counterparts in other societies (Collins 1979; Kerckhoff and Bell 1998), and where labor markets reward college attainment in stark contrast to how they reward high school achievement. In terms of economic returns to education, the wage premium for college attainment has been widely documented among scholars, is well known among lay observers, and has increased dramatically over the last several decades. Social scientists have also long argued that higher education, as an institution, is characterized by biased selection processes and that it stratifies life chances by systematically sorting people toward and away from valued social locations (e.g., Kerckhoff 1976; Gamoran 1992; Kerckhoff 1993; Natriello 1994; Kerckhoff and Glennie 1999; Loveless 1999; National Research Council 1999; Kerckhoff 2001, 3; Mullen 2010). In short, at various levels of analysis there is something about the system of formal higher education per se that affects people's life chances in profound ways.

In the context of the current work, we would expect someone who earned a four-year degree in engineering to have greater earning power than someone with a degree in general studies. Individual behaviors and personality characteristics only partially explain such differences in educational achievement (and thus socioeconomic status in adulthood), and we know even less about the reasons why some people finish their undergraduate educations earlier or later than others. Recent work in the status attainment, human capital, and stratification traditions has focused on the issue of timing in people's college experiences, mostly with regard to delayed college enrollments and transitions/transfers (e.g., Rowan-Kenyon 2007; Goldrick-Rab and Han 2011; Roksa and Velez 2012), but people's odds of college completion at discrete ages (in the context of "individual time" in individuals' lives) remains underexamined. The current work attempts to address some shortcomings in the fields of stratification and sociology of education, with specific research questions posed below.

Family background characteristics affect people's educational trajectories, yet there is still much uncertainty about specifically how this occurs (Blau 1999; Elman and O'Rand 2004; Rauscher 2011). From earlier work to recent scholarship, research has revealed some basic mechanisms by which family background may affect eventual educational attainment. Some argue that the most important dimensions of family background are genetic endowment and IQ (Jencks et al. 1979; Bowles and Gintis 2001), or characteristics of the home environment (Gottfried 1984), or family income (Hill and Duncan 1987). More recently, others have argued that family background is best represented by family forms and social psychological factors (Bowles, Gintis, and Groves 2008). In whichever ways family background may matter for educational attainment, specific questions about how it affects the timing of college completion become complicated by the very factors that comprise family background. Conceptually, family

background refers to dimensions like number of siblings and sibling achievements, parental education, parental relationship quality and intactness, transitions in family structure, home literacy environment, parenting styles and dispositions, socialization environment, neighborhood quality, peer group affiliations, and family socioeconomic characteristics such as wealth, income, parental education, and parental occupational stability.

In short, current evidence is not definitive regarding family background's effects on educational attainment. Though scholars have understandably hedged on proclaiming any single primary or "strongest" causal factor, a careful look at early and current research indicates some individual and family-level characteristics that may be more likely than others to affect the timing of people's educational attainment (Blau and Duncan 1967; Featherman and Hauser 1978; Blake 1981; Powell and Steelman 1993; Conley et al. 2003; Lareau 2003; Bowles, Gintis, and Groves 2008). Drawing on scholarship in education and socioeconomic stratification as well as life course theory, I decided to focus on gender, racial-ethnic identification, poverty in family-of-origin, mother's education, teenage parenthood, number of siblings, and eldest sibling's educational attainment.

Generally I ask which family background variables and individual characteristics help explain people's likelihoods of completing college at any given age. Specifically, and as a foreshadowing of the discussion and analyses to follow, three primary questions serve as a basis for the study. First, what are the time-related patterns of college completion among individuals who were entering adulthood in the late 1970s and early 1980s? Second, controlling for family socioeconomic conditions and other factors, what effect do minority racial-ethnic status and gender have on the time it takes people to complete college? Third, how might teenage parenthood and family poverty account for the apparent effect of minority race-ethnicity on

people's odds of college completion? Below, after describing the NLSY sample, I discuss the specific rationale for selecting the variables previously mentioned.

### **Data, Methods, and Measures**

The people who comprise the National Longitudinal Survey of Youth 1979 (NLSY) were born between 1957 and 1965.<sup>22</sup> The study sample comprises respondents from 26 waves of the NLSY, a nationally representative sample initially with 12,686 individuals who were ages 14 to 22 when enrolled in the survey beginning in 1979.<sup>23</sup> The NLSY is a multistage probability sample meant to represent the entire population of “youth” who resided in the United States in 1979. Sample respondents were re-interviewed annually from 1979 to 1994 and biennially thereafter (1996, 1998, and so forth), and I used data for the time period 1979 to 2004. The NLSY has advantages for an event history analysis because data about the timing of educational completion and many of its potential predictors were collected contemporaneously at each survey wave.<sup>24</sup> The final sample for this section's analyses comprised 4,421 respondents for whom there were complete records over the 25-year period 1979 to 2004.

---

<sup>22</sup> NLSY respondents are part of the “baby boomer” generation, or at least the young baby boom of people born between 1954 and 1965. Elman and O’Rand (2004, 132) have posited several reasons why baby boomers are appropriate subjects for analyzing the educational attainments of currently middle-aged adults. First, baby boomers were a large cohort who encountered labor markets in the 1970s and 1980s that had higher job skill requirements than in previous decades. This also means baby boomers were members of the generation that was able to reap the increased economic returns to those skills in the 1980s. Baby boomers also experienced the major economic downturns of the 1970s and 1980s, and they were the generation most involved with the historic increases in adult higher educational enrollments during those two decades. One basic but intriguing question is whether we should expect more or less heterogeneity in timely college completion among members of this cohort, an issue to which I return later.

<sup>23</sup> Please see Appendix C for a more detailed discussion about the nature of NLSY data.

<sup>24</sup> I excluded the NLSY's special military sample because the majority of these respondents were not interviewed after 1983. I also excluded non-interviews and observations with missing data, as has been done in other studies using NLSY data (e.g., see Klepinger et al. 1995; Garasky 1995). Finally, although it may seem unconventional, I did not use sampling weights. After consulting the NLSY technical documentation (NLSY 1983), I made this decision based on some comparisons of initial models with and without weights as well as personal correspondence with Dr. Jay Zagorsky of the Bureau of Labor Statistics, who is an expert in NLSY technical matters (Zagorsky 2010).

## Methods

“Survival” models, which are designed to make time-until-event the analytical focus, allow researchers using longitudinal data to offer more sophisticated treatments that account for the timing of social phenomena. Although I am most interested in the substantive uses of this approach, it also has important statistical advantages regarding issues such as error distributions and censoring. These issues are beyond the scope of the current work and have been well discussed in the methodological literature (e.g., Mayer and Tuma 1990; Singer and Willet 2003). Put simply, the event history approach has gained widespread popularity in mainstream sociology, and has been used in studies of various life events.<sup>25</sup> Considering how influential the life course perspective has been in various scholarly domains, and especially given its emphasis on the importance of timing in social stratification processes, one would expect greater amount of scholarship on timing of college completion than currently exists. With this in mind, an event history approach is a methodologically useful way to address the substantive issues of college completion from a life course perspective.

Here, I use a discrete time event history analysis to statistically model occurrence of college completion at any given age, with a model of the following form. :

$$\ln\left(\frac{P_{it}}{1-P_{it}}\right) = \alpha + f(t) + \sum_{k=1}^K \beta_k X_{ikt}$$

$P_{it}$  represents the hazard, that is, the probability (instantaneous “risk”) that the  $i^{th}$  respondent completes college at time  $t$  given that s/he has not yet completed.  $f(t)$  is some function of time that represents the baseline hazard, that is, the dependence of the hazard of college completion purely on time but regardless of the effect of any of the substantive X

---

<sup>25</sup> Such “life events” include but occupational mobility, onset of sexual activity, family formation and change, offspring educational achievement, as well as educational progressions, transfers, pathways, and so forth (e.g., Jacobs and King 2002; ; Bozick and DeLuca 2005; Bahr 2009; Wagmiller et al. 2010; Roksa and Velez 2012).

variables. As Singer and Willett (2003) suggest, the baseline hazard function can be thought of as analogous to the role of an intercept in ordinary regression. Here,  $\alpha$  is a constant,  $f(t)$  is a function of time,  $X_{ikt}$  is the value of the  $k^{th}$  covariate at time  $t$  for respondent  $i$ , and the  $\beta_k$  are coefficients associated with the various  $k$  explanatory or independent variables. The regression coefficients ( $\beta$ ) here indicate the change that can be expected in the odds of college completion, in relation to changes in the explanatory variables. In this basic model, the assumption is that a variable's effect (e.g., the effect of gender or race-ethnicity on the log-odds of college completion) remains the same at all points during the risk period. That is, *time* does not have interactive effects in combination with the various explanatory variables. The most direct interpretation of the estimates is in terms of exponentiated  $\beta$  coefficients, which measure the independent variables' multiplicative effects on odds of event occurrence at a given point in time (Box-Steffensmeier and Jones 2004).

In panel data sets like the one used here, data collection ends before all respondents have completed their educations (if in fact they would ever have completed had they been followed long enough). This phenomenon is termed “right censoring” in event history models, and applies to all NLSY respondents who had not completed a four-year degree by 2004<sup>26</sup> As discussed in the previous modeling comparison section (Section 2), with traditional regression methods two common practices have been to simply omit censored respondents or to impute values for the time-to-event variable. The problem is that this may introduce systematic error into the data. As Willett and Singer (1991, 409) have shown, “imputation tends to underestimate overall time in the risk set” because, for censored respondents, ultimate event-times are necessarily greater than the imputed values. In short, survival or event history analysis is said to handle the censoring

---

<sup>26</sup> “Right-censored” respondents are individuals for whom time-to-completion was unknown up until and including the final wave of data collection.

dilemma more evenhandedly because it incorporates every respondent's information until data collection ends.

### *Measures*

The outcome variable is "college completion," i.e., whether or not an individual reported having completed a four-year undergraduate degree at a given age. Time is a central variable here, measured as an individual's age in years. I followed respondents up to the point at which they completed four-year college or up to the most recent age for which information was available if they still had not completed. Below, in order to provide some context regarding the independent variables and why they were included, I will mix together some review of the literature as well as a discussion of measurement.

A key independent variable is gender, as women have long been more likely than men to complete high school and college (Goldschmidt and Wang 1999; U.S. Census Bureau 2010). I also included racial and ethnic status as an independent variable (black and Hispanic, with white as the reference category), which simply will be referred to as "race-ethnicity" throughout.<sup>27</sup> Black and Hispanic Americans typically have lower educational completion rates compared to whites, though this could be due to socioeconomic and family background factors associated with race-ethnicity (Hauser, Simmons, and Pager 2004). Studies that adjust for the differences in economic endowments that blacks and whites "bring with them" into educational systems have shown the black high school completion rate to be comparable to that of non-Hispanic whites

---

<sup>27</sup> The race-ethnicity measure for "black" is based on the NLSY interviewers' judgments rather than respondents' self-identifications. Appendix A (toward the end of this document) provides a fuller discussion of this decision about "race-ethnicity." Also, as has been done in most other studies using the NLSY 1979 cohort, "Asian" and "American Indian" respondents were omitted due to substantial *n* limitations in the initial sample.

(Conley 1999), but I am not aware of research that makes a similar argument about black-white differences in college completion.

A family's economic resources, by which I mean a measure of family poverty during respondents' mid- to late-adolescence (discussed below), is another key explanatory variable. Family economic resources are important because they can affect a person's educational attainment in complex ways (e.g., McLanahan and Sandefur 1994), and they provide children with various options including after-school education and other enriching or advantageous life experiences (Duncan 1994). There are many ways to conceptualize economic resources in a person's family of origin, and these include family poverty, parental earnings, income, wealth, and so forth. Parental financial resources have long been known to affect children's academic performance and educational attainment (McLanahan and Sandefur 1994; Duncan 1994). Compared to earnings, family income provides a more inclusive measure of a household's economic status. It is also more stable than earnings or wealth, thus I follow the lead of other researchers (Mayer and Lopoo 2004) in assuming that family income among children aged 14 to 21 is strongly correlated with what their family income was when they were very young (birth to early teens). The experience of poverty, as one indicator of family economic resources, influences a family's ability to provide or afford enriching after-school educational opportunities, which could plausibly affect later economic well-being in adulthood. Whether or not a person's family-of-origin experienced poverty for at least one year is commonly used as a relevant measure of family socioeconomic resources,<sup>28</sup> so I included a variable indicating whether

---

<sup>28</sup> This measure of family poverty may seem somewhat crude, but it differs from the "window variables" discussed in the dissertation's introduction (Section I) because families and individuals who had spent "at least one year below poverty line" typically would also qualify as having been two or three years below poverty line, a meaningful correlation that hints at a kind of convergent validity among the latter and former variables. By contrast, measurement validity is more pernicious with true window variables (e.g., the variable "lived with both biological parents at age 14" is not necessarily associated with the typical/modal type of parental presence during an individual's childhood years).



respondents lived below the federal poverty line (adjusted to the consumer price index for each year 1979 to 2004) for at least one year while living in their family-of-origin's household.

To further account for family socioeconomic status, I included a parental education variable for respondents' mothers' total years of formal schooling, a value that ranged from 4 to 21 years. In the NLSY, parental educational attainment appears to correlate strongly with offspring educational attainment as well as other offspring outcomes, and studies of "level of schooling" (not timely completion) have shown that each additional year of maternal education increases the likelihood that a child will complete high school (Garasky 1995).<sup>29</sup> Following other studies (e.g., Conley 2001), I include an indicator of the mother's number of years of education, where a value of 12 refers to high school degree, 16 indicates a four-year college degree, and 17 or greater indicates post-baccalaureate schooling. Based on conventional wisdom in social mobility research, I expect parental educational attainment, especially that of mothers because they are the most consistently present parental figure across household types, to strongly predict respondents' chances of timely college completion.

Additionally, there is evidence of a "correspondence" between an individual's educational attainment and that of their eldest sibling's education. Though not definitive, the evidence of cross-sibling influence in educational attainment has been documented mainly in terms of an older-to-younger sibling role modeling effect (Benin and Johnson 1984; Hauser and Wong 1989). A person's eldest sibling's education has been shown to be associated with several important later-life outcomes including educational attainment, so I included as well as eldest sibling's education in total years of formal schooling.

---

<sup>29</sup> These basic and unsurprising features of this dataset are consistent with much research in sociology and they also align with findings from much earlier status attainment models in mobility research, particularly those that focused on U.S. society.

The presence of siblings represents an important dimension of an individual's family context.<sup>30</sup> Parents' ability to provide attention and nurturing interactions with a child change (and tend to diminish) with an increase in the number of children in a household. In addition to sound theoretical reasons for a sibling effect, there is some evidence that number of siblings negatively influences children's educational attainment, in part because having more siblings can dilute family resources and cause relationship strain between parents and general stress within the household (Kuo and Hauser 1996). With this in mind, I included a count of siblings indicated by the number of siblings a respondent ever had from birth to age 18. Discerning how sibship figures into the timing of people's educational completions is complicated; as others have done, I simply assess whether having more siblings during childhood and adolescence may attenuate individuals' discrete-time risks of completing college (Blake 1989; Conley 2001).

Another explanatory variable, teenage childbearing, is defined as whether a respondent became a parent before turning 20 years old. Some researchers argue that, net of various background and personal characteristics, having a child during one's adolescent years does not significantly increase young mothers' risk of dropout from high school or their likelihood of college completion, and other scholars have suggested that unmeasured family background characteristics may attenuate the apparent disadvantages of teen childbearing (Hoffman, Foster, and Furstenberg 1993; Hoffman 1998).<sup>31</sup> There is a tremendous amount of scholarship on

---

<sup>30</sup> There are many dimensions of family context that one might include. To assess the influences of family structure on educational attainment, initially I focused on the *parental situation* in respondents' households during the formative years from birth to age 18, in particular the timing of and reasons for major transitions in respondents' family structures. These variables added minimal to no explanatory power, so I decided to drop them from the event history analyses, along with several other potential indicators of family context such as "home literacy" environment. Appendices B and C present some initial analyses and explain the reasoning behind these decisions.

<sup>31</sup> Personal characteristics other than teenage parenthood (but that may be closely associated with it) such as precocious sexual activity, parental characteristics, and adolescent "problem behaviors" negatively affect the likelihood of educational completion when researchers have controlled for the effect of teenage parenthood (Upchurch and McCarthy 1990; Francesconi 2007; Frisco 2008). In my discussion here, I simply mean to show that

teenage parenthood. In short, researchers, policymakers, and the lay public have long decried the apparent disadvantages of teenage childbearing, especially the presumed negative consequences for unmarried or unattached teenage women and eventually for their children (Moore and Waite 1977; Furstenburg et al. 1989; Furstenburg 2003). There has been vigorous debate about the presumed causal connections between teenage parenthood and socioeconomic disadvantage (see for example, Geronimus et al. 1999; Luker 1996; Lawler et al. 2001; Geronimus 2003), both of which may diminish a person's chances of ever completing college. As a general matter this debate has influenced my interest in the connection between age at college completion and teenage parenthood, but specifically I am most interested in whether teenage parenthood may mediate the apparent effects of racial minority status and sibship context on people's chances of college completion at any given age. Currently there is no social science consensus about the causal influence of teen parenthood on a person's chances of timely college completion or about its potential as a mediating influence between other characteristics and college completion. With this and the preceding discussions in mind, I pose and discuss three research questions addressed subsequently in the analyses.

## **Research Questions**

As mentioned previously, three primary questions serve as a basis for the study. The first question is: *What are the time-related patterns of college completion among individuals who were entering adulthood in the late 1970s and early 1980s?* The second third of the twentieth century was a period of unprecedented increases in public funding for and interest in higher education; the U.S. economy was characterized by periods of growth and stagnation, and still

---

the aforementioned conventional (and often ideological) assumptions about teenage parenthood's apparent negative effect on people's lives have invoked empirical questions that remain unsettled in the sociological literature.

had a strong manufacturing industrial base; and college was coming to be seen as possible for more than just the elite as it had been thought of in previous eras. Considering this social context in which young baby boomers came of age during the late 1970s to the early 1980s, I expect the majority of people in the NLSY to have highest odds of completion during their very early twenties, with much lower odds of completion after the age of 23 or 24 (in most studies of educational attainment, ages 24 and 25 are the conventional cutoff points for defining “timely” college completion).

A second question is: *What effect do minority racial-ethnic status and gender have on the odds of individuals’ college completion at given ages, controlling for family socioeconomic conditions and other factors?* Based on theory and previous scholarship, I expect racial-ethnic minority status will dampen people’s odds of completion at any given age and also that men will be disadvantaged compared to women on this count.

A third question related to the second is: *To what extent do teenage parenthood, family poverty, and maternal education account for the expected negative effect of minority race-ethnicity on the odds of people’s college completions at any given age?* For various reasons discussed later in the subsection on decomposition models, I expect intervening effects of teenage parenthood and family poverty will be more substantial than maternal education, and that the degree to which these variables account for the minority-college association will be more substantial for blacks compared to whites and Hispanics.

### **Sample Statistics and Regression Models for College Completion**

Table 3.1 displays selected descriptive statistics regarding individuals for whom there was valid information on explanatory variables during the 25-year observation period (between 1979 and

2004). The percentage who completed college at some point in their lives (summarized in the “ever completed” column) is only slightly higher than the percentage of similar-aged people living in the U.S. who had completed college during the late 1970s and into 1980s and 1990s; among the sample respondents 19.4 percent completed college at some point during the observation period.<sup>32</sup> The proportion of women who completed was slightly greater than among men (52.2 percent and 47.8 percent, respectively), while the proportion of non-Hispanic whites who completed (23.6 percent) was nearly twice that of Hispanics (11.3 percent) and nearly nine percentage points greater than that of Blacks (14.7 percent).

Individuals whose families of origin experienced at least one year in poverty had substantially lower proportions of college completion (9.7 percent) compared to those who grew up in financially better off families (25.9 percent). There are many mechanisms by which we should expect family economic advantage to lead to higher odds of educational attainment, including the fact that children from better off families benefit from their parents’ resources (material and otherwise) and they tend to have higher educational aspirations and academic motivations (e.g., Coleman 1988; Hauser 1980).

Also unsurprising was the finding that individuals who became parents during their teenage years had a dramatically lower completion rate (2.9 percent) relative to those who did not become teen parents (22.9 percent). Parenthood can seriously divert a person’s material resources, time, emotional energy, and motivation away from schooling and toward other priorities. As well, for many reasons it could be that those individuals who become teen parents were already less likely to ever enroll in college. Whatever the mechanism, researchers have

---

<sup>32</sup> As mentioned earlier, in a span of just fifty years not only did high school completion become normative among the general population (with 34 percent of Americans having completed high school in 1950 compared to 80 percent in 2000), but the share of American adults who had completed college *quintupled* over those same decades, from just over 5 percent to 25 percent as of 2010 (U.S. Census Bureau 2010).

exhaustively shown that adolescent entrance into some normatively “adult” roles unrelated to education, such as marriage and parenthood, portends lower overall educational attainment.

Among individuals whose mothers had at least 12 years of formal education, 29.4 percent completed college, nearly four times that of those whose mothers had less than 12 years. This was expected, as parent-child correlations of education have long been quite strong in the United States. There also is ample evidence over many decades that shows strong educational resemblance among siblings, so it was no surprise to see that 23 percent of individuals whose eldest sibling had attained at least 12 years of schooling ended up completing college, more than double the percentage of those whose eldest sibling had attained less than 12 years. And finally for this “ever completed” indicator, 31 percent of individuals who had fewer than three siblings eventually completed college, more than twice that of those who grew up with three or more siblings. One highly plausible mechanism for this is that additional siblings within a family can have negative effects on the resources (time, money, energy, etc.) available to each child, and there are many reasons why the presence of more siblings could attenuate a person’s odds of finishing a four-year degree.

The rightmost column in Table 3.1 summarizes timing of college completion, with median age at completion shown for each of the preceding variables. Patterns of differences go in theoretically expected directions. Median age at completion was one year older among individuals whose mothers had less than 12 years of education as compared to those whose mothers had at least 12 years of education; older median age at completion was also observed for individuals who had three or more siblings compared to those who had fewer than three. Such basic patterns also pertain to blacks and Hispanics as compared to white respondents, to men

compared to women, to teen parents as compared to non-teen parents, and so on down the left hand list of variables.

**Table 3.1**  
**College Completion Statistics for Selected Variables**

Variables	Ever completed college (%)	Median age at college completion (min/max, s.d.)
<b>Gender</b>		
Men (n =2,229)	410 (47.8)	23 (20/41, 4.1)
Women (n =2,192)	447 (52.2)	22 (20/43, 4.9)
<b>Race-ethnicity</b>		
White (n=2,611)	616 (23.6)	22 (20/43, 4.0)
Black (n=1,083)	159 (14.7)	23 (21/43, 5.5)
Hispanic (n=727)	82 (11.3)	23 (20/42, 5.2)
<b>Teen Parenthood</b>		
Yes (n=784)	23 (2.9)	31 (21/43, 7.5)
No (n=3,637)	834 (22.9)	22 (20/42, 4.2)
<b>Family Poverty</b>		
Yes (n=1,781)	172 (9.7)	23 (21/43, 5.6)
<b>Table 1 (continued)</b>		
No (2,640)	685 (25.9)	22 (20/43, 4.1)
<b>Mother's Education (years)</b>		
≥ 12 yrs. (n=2,379)	700 (29.4)	22 (20/43, 4.0)

< 12 yrs. (n=2,042)	157 (7.7)	23 (21/43, 5.9)
<b>Eldest Sibling's Education (years)</b>		
≥ 12 yrs. (n=2,958)	683 (23.1)	22 (20/43, 4.5)
<12 yrs. (n=1,463)	174 (11.9)	23 (21/43, 4.6)
<b>Number of Siblings</b>		
≥ 3 (n=3,215)	483 (15.0)	23 (21/43, 5.0)
< 3 (n=1,206)	374 (31.0)	22 (20/42, 3.6)
<b>N = 4,421</b>	857 (19.4)	23 (20/43, 4.5)

In other descriptive results not shown, 27.7 percent of Black respondents became parents during their teenage years as compared to 21 percent of Hispanics and only 13.8 percent of whites. A slight majority of Black and Hispanic respondents' families of origin experienced at least one year in poverty (56.6 percent and 50.1 percent, respectively), nearly double what the percentage was for whites (28.1 percent). The same differences by race-ethnicity in median age at college completion pertained among those who lived in poverty and became parents as teenagers. That is, 22 years was still the median age at college completion among whites who lived in poverty and were teenage parents, and 23 years was the median age at college completion for Blacks and Hispanics whose families of origin experienced poverty and who were teenage parents (identical to the median ages reported in the far right column of Table 3.1). As expected, there was an increase in the *mean* age of college completion, with the rises in mean age going in expected directions for members of each racial group (mean age was 3 years greater for Blacks who met these conditions, 3 years greater for Hispanics, and 2 years greater for whites).



The foregoing descriptive view should not seem surprising vis a vis conventional wisdom in social science, and it represents a starting point from which I attempted to model potential connections between the various predictors and individuals' odds of completing college. But before discussing findings from the full regression models, it is useful to display in greater detail the distribution of college completion across ages among respondents in the sample; this implies the most basic regression, namely people's chances of college completion regressed on yearly age intervals.

### *The Discrete-time Hazard Rate for College Completion*

Figure 3.1 summarizes the changes in hazard rate for college completion over the entire risk period up to age 47. The hazard function represents the conditional probability of completing college at each age among respondents who had not completed at earlier ages; the greater the hazard in a given time period indicates a greater estimated "risk" that eligible respondents would have of completing college at a given age.<sup>33</sup> The period of highest risk of college completion among respondents who finished college at some point (or were censored) is the early adult years, with individuals' risks of completion decreasing sharply thereafter (but with a subtle wrinkle that will be discussed below).

Put another way, Figure 3.1 represents the relationship between age, or "analysis time" on the x-axis, and yearly risk of college completion, or hazard rate, shown on the y-axis.

Expected ages for risk of completing college are shown at the bottom of the figure: Ages 21 to 23 were the periods of substantial risk for college completion among all individuals in the

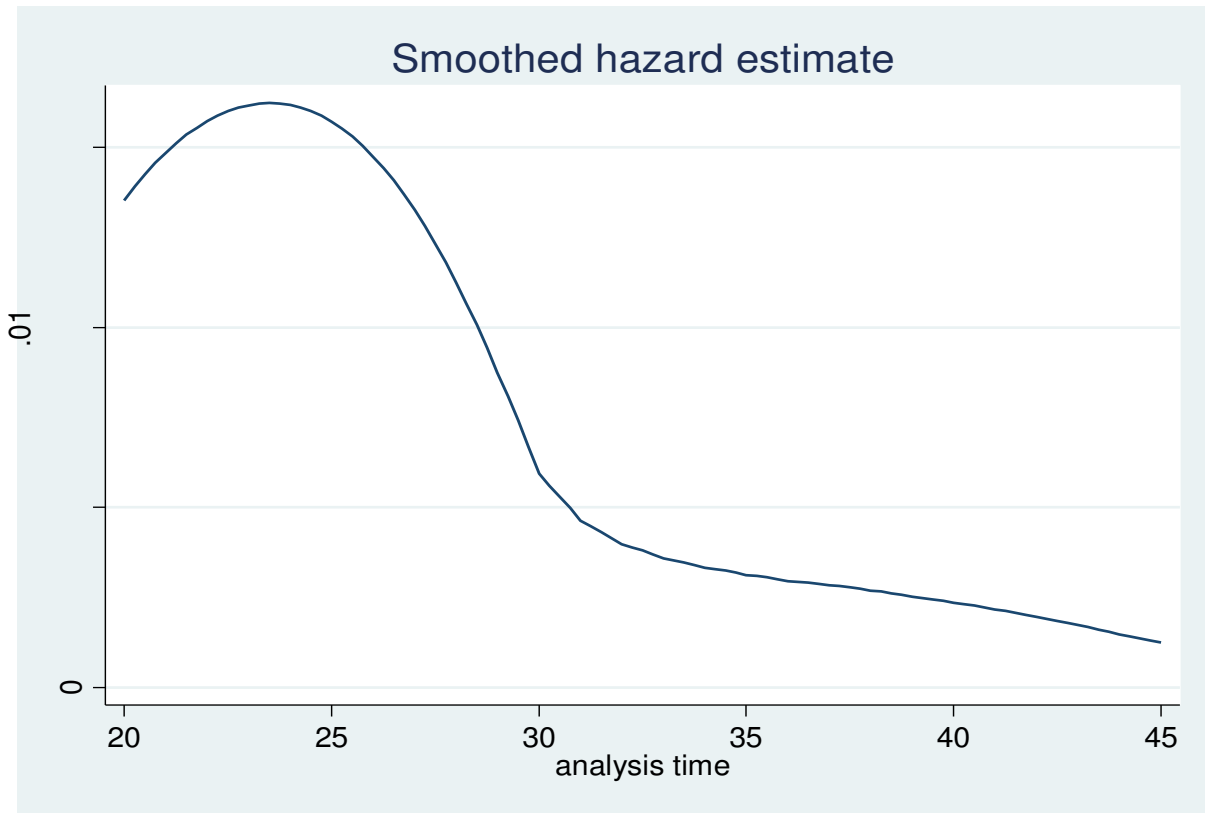
---

<sup>33</sup> By contrast, the survivor function (as discussed in Section 2) cumulates the period-by-period proportion of individuals who remained in the risk set by given ages, i.e., the proportion of respondents who had not yet completed college at each age interval. Put another way, the proportion of individuals who remained eligible for completing college began to drop beyond age 24.

sample (age 22 being the age of greatest risk), there was still substantial risk for completion between 24 and 26, and then risk of completion dropped off considerably beyond age 26 with a less precipitous decline continuing beyond age 30. In the language of survival analysis, compared to earlier ages there is a much lower likelihood of people ever completing college after they reach age 30 if they had not yet completed by then. Individuals who had not yet completed college at each age period were most at risk of completing prior to age 26, with 20 to 26 generally being the period of greatest risk. The hazard function peaks during respondents' early to mid-twenties and declines continuously thereafter.<sup>34</sup> Although the majority of individuals who ever completed college between 1979 and 2004 did so before turning 30, there was still a notable amount of "remaining risk" as individuals aged into their later twenties (26, 27, etc.).

---

<sup>34</sup> Unlike the survivor function, hazard estimates can increase or decrease in value over time; the fact that hazard never increases beyond age 20 reflects the age-graded nature of college completion in U.S. society (though, as we saw in Section 2, college completion is not as rigidly age-graded as high school completion).



**Figure 3.1**  
**Estimate of Hazard Rate (y-axis) for College Completion by age (x-axis)**

*Life Course Theory and Heterogeneity in the Timing of College Completion*

According to life course theory and research (e.g., Pallas 1993; Bozick and DeLuca 2005), in general we should expect this observed risk of college completion to decline as individuals enter their early mid-twenties, and many explanations would fit the general story portrayed in Figure 3.1. First, individuals who enroll in college at age 18 or 19 could experience life course transitions such as parenthood or employment that may diminish their chances of on-time completion. Second, people who were initially off-time in college enrollment may never complete college for many reasons, perhaps because those who did enroll in college during their mid-twenties may have found themselves competing with younger and more adept peers, which

could negatively influence their willingness or interest in staying in school. As well, individuals who enrolled (or who might have considered enrolling) during their late 20s or early 30s may have had family and work commitments that clashed with college-related obligations. In short, there were no surprises in this regard.

However, even with the sharp decline in individuals' chances of completion, Figure 3.1 above shows that there was still substantial chance of completing college among all respondents as they aged into their mid-20s (ages 24 to 27). Torche (2011, 767), in her examination of cohort trends for age-adjusted college completion rates among people born between 1905 and 1965, identified a "period expansion" in college completion for those born between the 1910s and the late 1940s, a "period decline" in completion for those born in the 1950s, and another expansion for late Baby Boomers born in the 1960s. Reasons for the expansions and declines in college completion include growth in earnings returns to schooling, variation in federal investments toward higher education, legalized racial segregation, the emergence of the G.I. Bill, and sudden availability of college draft deferments. It was among the generations that came of age during the first two-thirds of the twentieth century that college completion became established as "age-normative" around the 20- to 22-year old mark. Respondents in my sample were born during the late Baby Boom (ending in 1965) which was the tail end of the aforementioned period, thus an important descriptive matter is how much heterogeneity in age-at-completion we should expect among them.

Considering that the rise of a substantial non-traditional aged student population did not begin until late in the twentieth century, one would expect relatively little heterogeneity among late Baby Boomers, i.e., that they would have typical ages at completion similar to those of earlier generations. But the curve shown in Figure 3.1 suggests a somewhat unexpected

heterogeneity in the timing of people's risk of college completions, which is intriguing because it suggests that post-secondary education among members of this baby boom cohort was not as rigidly age-graded as casual observers might expect. This result fits with one of the most consistent findings in life course research since the 1980s, namely that a modal sequence of life events does seem to exist among the population—as Hogan (1980) noted long ago, from about 1940 to the early 1970s the most commonly ordered path was schooling, employment, then marriage—but the proportions of people following this normative order has *decreased* over the decades, especially during the last quarter of the twentieth century and into the current one (Marini 1987; Hogan and Mochizuki 1988; Modell and Goodman 1990; Shanahan 2000). Many studies have shown that, since approximately the 1960s, interruptions in schooling are a more common feature in the transition to adulthood than they were during the first half of the twentieth century, and we may be observing the effects of this society-wide shift with the result presented in Figure 3.1. (People were completing college later in their twenties, but with the NLSY's relatively limited 7-year age range between cohorts it was difficult to discern this graphically. When I graphed the hazard for the different cohorts within the NLSY—comparing individuals who were 15 or 16 in 1979 with those who were 19, 20, and so on—the changing of the aforementioned pattern *can* be seen in the difference between 15 year-olds and 19 year-olds but *not* in the differences between other cohorts.) In short, interruptions often turn into *disruptions* that put people off-track from college permanently. The basic descriptive finding displayed here implies that, even among members of an earlier generation for which one might assume mostly conventional or “traditional” ages for college completion, it seems we should not claim that the early twenties was the age period exclusively associated with college completion.

### *Brief Note on Modeling Discrete-time Risk of College Completion with Logistic Regression*

The descriptive results presented above represent a starting point from which we can examine regression models of the various predictors' effects (direct and indirect) on conditional odds of completing college. Here, I present results from several discrete-time survival models, having applied logistic regression to a person-period (yearly) data file. Subsequent discussions will consider how teenage childbearing, family poverty, mother's education, and other variables may help elaborate our understanding of race-ethnicity and gender as predictors of college completion.<sup>35</sup> A first step in this analysis, though, was to develop a time-only model of college completion. (While not of substantive interest, this is a necessary and standard feature in any discrete time survival model.) In brief, models not shown here indicated that the general time specification should include every possible age as a predictor—14, 15, 16, and so forth, up to age 47.<sup>36</sup> I decided that the most convenient and parsimonious specification was one by which the “risk period” begins at age 20, basically the youngest age at which individuals completed college, and that it should end with either the completion of college, or age 46, or the point at which an individual stopped contributing information to the dataset (the point at which they were right-censored). This age specification represents the general context for regression models of odds of college completion that I will now discuss.

---

<sup>35</sup> Coefficients are shown as odds ratios, which in logistic regression represent the exponential of the log-odds ratio. This statistic provides one way to address questions of the form “How does a unit-change in an independent variable alter the predicted odds of some event occurring?” As has been well discussed in the methodological literature (Menard 2002, 57), the odds ratio is unique in how it presents information. Unlike probability, the odds ratio can range from 0 to infinity, with 1 indicating no relationship, and less than 1 or greater than 1 indicating negative or positive relationships. For example, an odds ratio with a value greater than 1 means that the discrete-time odds of completing college increased with an increase in the value of an independent variable, and an odds ratio of less than 1 means that the odds of completing college at given ages decreased with an increase in the value of an independent variable. For categorical independent variables, odds ratios are interpreted as between groups (e.g., men compared to women, individuals who became parents as teenagers compared to those who did not). For continuous independent variables, odds ratios are interpreted as between those individuals who are identical on other independent variables but who differ by one unit on the variable of interest.

<sup>36</sup> Among these possibilities, I compared goodness-of-fit—using many discrete ages at first, and then using various age-range categories—in order to discern the optimal modeling form to represent “time-to-college-completion.”

## Regression Models for Conditional Odds of College Completion

With an outcome like college completion at any given age, many differences are worth analyzing. I begin by noting that individuals from economically disadvantaged families, racial-ethnic minorities, and men were less likely to ever complete college compared to those who grew up in economically advantaged families, whites, and women. In the NLSY sample examined here we will see that the influences of economic advantage and gender seem very robust, and that racial-ethnic differences do appear when we model discrete-time odds of completion but these differences narrow when other factors are taken into account. (As a brief reminder that I will reiterate on occasion: While I will tend to use the phrases “odds of completion”, “odds ratio,” and “chances of completion” to describe results, these terms may be misleading and are just shorthand for the discrete-time outcome. That is, in hazard modeling scenarios, what I will refer to as the “odds of college completion” actually reflects a comparison of the *relative* odds for completion *at each age* across individuals’ lives. Technically speaking this refers to “odds of completion at any point in time given that they had not yet completed up to that point.”)

Parental financial resources can affect children’s academic performance and educational attainment (e.g., see McLanahan and Sandefur 1994 and Duncan 1994), so we might reasonably expect family economic disadvantage to be associated with lower odds of completing college at any age. As discussed previously, I followed other scholars (Mayer and Lopoo 2004) in using a measure of family poverty, indicated by whether an individual’s family was below the federal poverty line (adjusted to the consumer price index for the years 1979 to 2004) for at least one year while they were growing up. As expected, the results in Tables 3.2a and 3.2b (along with descriptive results displayed previously) show that the experience of family poverty is negatively associated with people’s odds of college completion at any age. Of course, the pathways

through which this effect operates could be related to various unobserved features of people's lives. Students whose families were socioeconomically disadvantaged may be less likely to enroll in college because their high school counselors did not see them as college ready and treated them as such, or because few of their peers had college aspirations, or because their parents lacked information about the application process and/or discouraged them from attending college. It could also be that socioeconomically disadvantaged students were forced (for financial reasons) to enroll in college at later ages, which would portend higher likelihood of eventual dropout.

Results from model 1 in Table 3.2a show that racial-ethnic minority status seemed to reduce people's odds of college completion as compared to whites, with blacks having less than two-thirds the odds (0.573) and Hispanics less than half the odds (0.477) of college completion at any age as compared to non-Hispanic whites. However, this apparent disadvantage for minorities diminished substantially as socioeconomic and family background controls were introduced in subsequent models. Specifically, accounting for teenage parenthood in model 2 considerably narrowed the black/white difference ( $OR = 0.664$  vs. 0.573 in the previous model.). Looking to model 3, which included family socioeconomic disadvantage during the years when respondents were living with their families of origin, we see that black respondents' odds were even closer to whites ( $OR = 0.809$ ). Differences between Hispanics and whites also narrowed across the models in similar direction ( $OR = 0.447$  in model vs. 0.806 in model 4) when these variables were added.



**Table 3.2a**  
**Discrete Time Survival Model of College Completion using Logistic Regression<sup>37</sup>**

Variables	Model 1 <i>OR</i>	Model 2 <i>OR</i>	Model 3 <i>OR</i>	Model 4 <i>OR</i>
<b>Black</b>	0.573** (0.509 - 0.643)	0.664** (0.590 - 0.747)	0.809** (0.696 - 0.941)	0.849* (0.728 - 0.989)
<b>Hispanic</b>	0.477** (0.412 - 0.551)	0.515** (0.445 - 0.596)	0.524** (0.431 - 0.636)	0.806* (0.659 - 0.987)
<b>Female</b>	1.140* (1.045 - 1.244)	1.338** (1.226 - 1.461)	1.334** (1.190 - 1.496)	1.303** (1.159 - 1.464)
<b>Teen parent</b>		0.129** (0.100 - 0.167)	0.137** (0.098 - 0.193)	0.168** (0.119 - 0.238)
<b>Fam. Poverty</b>			0.447** (0.386 - 0.516)	0.591** (0.509 - 0.686)
<b>Mother's edu.</b>				1.269** (1.239 - 1.301)
Constant	0.117** (0.107 - 0.127)	0.123** (0.113 - 0.134)	0.155** (0.138 - 0.173)	0.008** (0.006 - 0.011)
Observations	236,681	236,681	142,090	137,720
BIC'	-5389.473	-5853.947	-3654.877	-3958.951
Craig-Uhler R2	0.233	0.252	0.270	0.298
McKelvey/Zavoina's R2	0.631	0.653	0.707	0.715

odds ratios  
(95% confidence interval)  
\*\* p-value < 0.001  
\*p-value < 0.05

Turning now to Table 3.2b, we see that among Hispanics in the sample teenage parenthood and family poverty status do not seem to matter nearly as much as mother's and eldest sibling's education. The odds of discrete-time college completion increased most

<sup>37</sup> The table entries represent estimated conditional odds on completing college at any age, for individuals who had not yet completed.

substantially for Hispanics as compared to blacks when mother's years of education was added in model 4, a 54 percent rise in the odds ratio from the previous model (0.524 to 0.806). As well, the odds of completion among Hispanics went from (0.815 to 0.927) when eldest sibling's years of education was introduced in Model 6.

**Table 3.2b**  
**Successive Discrete Time Survival Models for College Completion using Logistic Regression**

<b>Variables</b>	<b>Model 4 OR</b>	<b>Model 5 OR</b>	<b>Model 6 OR</b>
<b>Black</b>	0.849* (0.728 - 0.989)	.0889 (0.762 – 1.037)	0.876 (0.726 – 1.056)
<b>Hispanic</b>	0.806* (0.659 - 0.987)	0.815* (0.666 – 0.998)	0.927 (0.727 – 1.182)
<b>Female</b>	1.303** (1.159 - 1.464)	1.301** (1.157 – 1.462)	1.291** (1.121 – 1.487)
<b>Teen parent</b>	0.168** (0.119 - 0.238)	0.168** (0.118 – 0.238)	0.167** (0.109 – 0.254)
<b>Fam. Poverty</b>	0.591** (0.509 - 0.686)	0.626** (0.538 – 0.728)	0.652** (0.544 – 0.782)
<b>Mother's edu.</b>	1.269** (1.239 – 1.301)	1.256** (1.224 – 1.287)	1.242** (1.206 – 1.279)
<b>Siblings</b>		0.929** 0.899 – 0.958)	0.912** (0.878 – 0.948)
<b>Sibling edu.</b>			1.042** (1.029 – 1.056)
<b>Constant</b>	0.007** (0.006 - 0.011)	0.011** (0.008 – 0.016)	0.008** (0.005 – 0.013)
Observations			

	137,720	137,631	100,796
BIC'	-3958.951	-3968.068	-2673.882
Craig-Uhler R2	0.298	0.300	0.297
McKelvey/Zavoina's R2	0.715	0.717	0.696

odds ratios  
(95% confidence interval)  
\*\* p-value < 0.001  
\* p-value < 0.05

### *The Influence of Eldest Sibling's Educational Attainment and Number of Siblings*

Since the early work of Hauser and Featherman (1976; also see Featherman and Hauser 1978), much evidence has corroborated their initial finding of a strong sibling resemblance in eventual educational outcomes. This, along with the enduring connection between parental educational attainment and offspring education, implies that the family is an important stratifying agent in U.S. society. In families where the eldest child has relatively low educational attainment, any given sibling's education level is likely to be relatively low, and conversely for families in which the eldest child has relatively high educational attainment. This may work through role modeling or other social psychological processes related to sibling influence. Whatever the mechanism may be, my findings align with the general received wisdom among stratification scholars—sibling's education should favorably influence people's chances of completing college—but not in nearly as pronounced a way as expected (in the final model, *OR* for eldest sibling's education was only 1.042). But while eldest sibling's education *per se* may not seem positively influential, its inclusion in the model was associated with a change in the effect of being Hispanic on people's odds of completion; the coefficient for Hispanic ethnicity went from an *OR* of 0.815 to 0.927 when eldest sibling's education was included in model 6 of Table 3.2b. (A brief reminder I mentioned earlier: While I use the phrases “odds of completion”, “odds ratio,” and “chances of completion” to describe my results, these terms are shorthand for the discrete-time outcome. In a hazard model like this, what I refer to as the “odds of college completion” reflects a comparison

of the relative chances for completion at each age across individuals' lives.)

The number of siblings in a family represents another important dimension of sibship, and it was included in the regression models displayed in Tables 3.2a and 3.2b. For theoretical reasons discussed earlier, I expected the presence of more siblings to substantially reduce people's conditional odds of college completion; initial descriptive results presented earlier in Table 3.1 showed that twice the percentage of individuals who had less than three siblings eventually completed college compared to those who grew up with three or more siblings. As mentioned in that subsection, there is ample research to suggest that additional siblings within a family can dilute the various resources (time, money, energy, etc.) available to each child, so there are good reasons to expect that the presence of more siblings will reduce people's odds of completing a four-year degree. However, regression results in Table 3.2b indicate only a slight negative association between number of siblings and completion ( $OR = 0.929$  and  $0.912$  in the final two models).

#### *Conditional Odds of College Completion: Separate Models by Gender and Race-Ethnicity*

Results from the initial regression models presented above indicated that odds of college completion at any age were lower for men, racial-ethnic minorities, individuals who were teen parents or whose mothers had lower levels of education, and those who experienced family poverty during their childhoods. The influence of the two sibling variables was smaller than expected; the coefficient for eldest sibling's education was smaller than what one would assume based on existing theory and research, and having more siblings seemed to have only a slightly negative effect on people's conditional odds for college completion. We will see that this finding (smaller than expected "sibling effects") also pertains to other models discussed below. Of

particular note from the results discussed previously was that controlling for economic disadvantage in individuals' families of origin (most pronounced among blacks) and for family members' educational background (most pronounced among Hispanics) seems to attenuate what first appeared to be a substantial negative effect of minority status on odds of college completion at any given age. This being said, when all controls were included in the full model, racial-ethnic minority status still predicted lower conditional odds of completing college as compared to whites. The possibility that the influence of minority status on completion may be (partially or perhaps mostly) mediated through other variables will be examined later with the decomposition models. For now we turn to regression models that display conditional odds of completion separately by gender and race-ethnicity, with various controls.

First it must be noted that, in analyses not shown, I examined models allowing for teen parenthood to have an interactive effect with gender and race-ethnicity, on the presumption that teen parenthood would, for example, be associated with lower conditional odds of college completion for females as compared to males. These interaction terms were added to the discrete-time model for college completion. Women who were teen parents had only slightly lower odds of completion at any given age than men who were teen parents, and black women who were teen parents had very similar odds of completion as comparable black men. Women who experienced family economic disadvantage were more likely to complete than men. Blacks and Hispanics who experienced family disadvantage had one-third and twice the odds, respectively, of completing compared to their white counterparts. However, the interactions were not statistically significant even at the 0.10 level (a lone exception to this was the interaction

term for family poverty and Hispanic ethnicity), so instead I estimated separate models by gender and race-ethnicity to do the group comparisons.<sup>38</sup>

We begin with a look at differences in the influence of certain variables on men in the sample as compared to women. In line with most studies of educational attainment, women had the edge in degree completion over time; we already saw in Tables 3.2a and 3.2b that women had higher odds than men at any age of completing college, regardless of control variables. (The odds at any age of college completion for women ranged from 1.14 to 1.38 times higher than for men.) Table 3.3 below displays more detail with separate regression results by gender, with two findings of particular note. First, among women with the controls included in Models 2 and 3, becoming a parent in adolescence was associated with slightly lower odds of completion than it was for men (in Model 3, *OR* = 0.157 for women vs 0.178 for men). The difference in teen parenthood's effect for men compared to women was surprisingly small. I expected that teenage parenthood would have a more substantially suppressive influence on women's chances of completion, at least in part because of women's greater childrearing responsibilities, which could steer women away from or otherwise take precedence over college-related obligations. Though this modeling scenario does not provide a definitive comparison, it does suggest that there are other conditions, events, and behaviors pertinent to the lives of people who become teen parents that, irrespective of gender, might lower their likelihoods of completing college (and that would also presumably put them "on track" for other events such as early marriage or entry into the workforce). This basic finding resonates with a vast scholarship in sociology because it suggests

---

<sup>38</sup> I recognize that estimating separate models by groups, instead of a single model with interaction terms, can create problems such as heteroscedasticity and a loss of statistical power. There also is a danger of claiming effect differences between groups even though such differences may not be statistically significant. (Interaction terms should more accurately indicate the difference in effects between groups.) However, there may be less reason for concern if the sample size for one group is close to that for the other group. It is with this in mind that I assume greater accuracy for estimates of effect differences by gender than for the other group comparisons, since the sample sizes were very similar for the men-women models but there were much bigger differences in sample size for the models comparing whites, blacks, and Hispanics.

that, though there may be something about experiencing teenage parenthood *per se* that alters people's life paths, any number of unobserved factors are also likely at play.

Mother's education was associated almost identically with higher odds of college completion for men and women, but this was not true regarding family poverty's influence. Among men, growing up in a family that experienced a year or more of poverty during childhood portended lower chances of college completion over time than it did for women (shown by Model 3 in Table 3 below,  $OR = 0.640$  for men vs.  $0.720$  for women). Though the experience of family poverty during childhood was associated with substantially lower odds for both men and women, it seems to harm women's chances a little less, at least as indicated by this set of comparisons.

The difference between men and women in the degree to which poverty in one's family-of-origin was associated with their odds of college completion was smaller in magnitude compared to the gender difference in the effect of eldest sibling's education. A one-year increase in eldest sibling's education was associated with an odds ratio of 1.17 in men's chances of college completion, approximately 14 percent greater than the degree to which sibling's education was associated with women's odds (1.03). When mother's education and the two sibling variables were included in Model 3, the odds ratio for family poverty's effect on college completion increased similarly among women and men. There was a notable difference in how the presence of more siblings influenced men's and women's chance of college completion over time; an additional sibling had a more suppressive influence on men's odds of completion (shown by Model 3 in Table 3.3,  $OR = 0.863$  for men vs.  $0.948$  women).

**Table 3.3**  
**Conditional Odds of College Completion Over Time, by Gender**

	Model 1	Model 2	Model 3
<b><u>Women</u></b>			
<b>Black</b>	0.552** (0.432 - 0.707)	0.784** (0.608 - 1.011)	0.857** (0.662 – 1.101)
<b>Hispanic</b>	0.344** (0.244 - 0.486)	0.443** (0.312 - 0.628)	0.778** (0.543 – 1.112)
<b>Teen parent</b>		0.124** (0.076 - 0.202)	0.157 (0.096 – 0.257)
<b>Fam. Poverty</b>		0.515** (0.406 – 0.654)	0.720** (0.561 – 0.923)
<b>Mother's Edu.</b>			1.229** (1.179 – 1.281)
<b>Sibling Edu.</b>			1.030** (1.013 – 1.047)
<b>Siblings</b>			0.948* (0.900 – 0.998)
<b><u>Men</u></b>			
<b>Black</b>	0.534** (0.413 - 0.691)	0.748** (0.574 - 0.975)	0.898 (0.684 – 1.181)
<b>Hispanic</b>	0.480* (0.349 - 0.662)	0.621* (0.449 - 0.858)	1.084 (0.777 – 1.514)
<b>Teen parent</b>		0.145** (0.064 – 0.326)	0.178** (0.078 – 0.400)
<b>Fam. Poverty</b>		0.383** (0.296 – 0.496)	0.640** (0.489 – 0.836)
<b>Mother's Edu.</b>			1.226** (1.175 – 1.281)
<b>Sibling Edu.</b>			1.169** (1.114 – 1.227)
<b>Siblings</b>			0.863** (0.814 – 0.916)
odds ratios (95% confidence interval) ** p-value < 0.001 *p-value < 0.05			



Previously it was shown that family poverty is negatively associated with anyone's conditional odds of college completion, and here we see some racial-ethnic differences in this association. As presented in Table 3.4, for each racial-ethnic group, family poverty predicted lower odds compared to those who had more advantaged family backgrounds, but it was associated with much lower odds of completion among whites and blacks than among Hispanics. As well, when mother's education and the two sibling variables were included in Model 3, the odds ratio for family poverty's effect on college completion increased much more among Hispanics than it did for blacks and whites (among Hispanics, family poverty's negative association with odds of completion basically disappeared in Model 3, to 0.996).

Returning to the issue of teenage parenthood, we see in Table 3.4 that for each racial-ethnic group it was negatively associated with odds of completion, with relatively small differences between groups (in Model 3,  $OR = 0.157$  for blacks vs.  $0.162$  for whites and  $0.202$  for among Hispanics). For most of the twentieth century the teenage birth rate was considerably higher among blacks compared to Hispanics and non-Hispanic whites in the United States (Martin et al. 2015).

**Table 3.4**  
**Discrete-Time Odds of College Completion, by Race-ethnicity-Ethnicity**

	Model 1	Model 2	Model 3
<b><u>Black</u></b>			
<b>female</b>	1.183 (0.863 - 1.621)	1.461* (1.063 - 2.001)	1.435* (1.042 – 1.976)
<b>Teen parent</b>		0.139** (0.068 - 0.285)	0.157** (0.076 – 0.322)
<b>Fam. Poverty</b>		0.507 (0.366 - 0.703)	0.729 (0.513 – 1.037)
<b>Mother's Edu.</b>			1.167** (1.088 – 1.252)
<b>Sibling Edu.</b>			1.123* (1.033 – 1.221)
<b>Siblings</b>			0.946 (0.883 – 1.013)
<b><u>White</u></b>			
<b>female</b>	1.138 (0.965 - 1.341)	1.319** (1.117 - 1.557)	1.340 (1.130 – 1.589)
<b>Teen parent</b>		0.113** (0.062 – 0.209)	0.162** (0.088 – 0.297)
<b>Fam. Poverty</b>		0.399** (0.315 – 0.507)	0.565** (0.442 – 0.723)
<b>Mother's Edu.</b>			1.315** (1.266 – 1.366)
<b>Sibling Edu.</b>			1.038** (1.023 – 1.054)
<b>Siblings</b>			0.859** (0.813 – 0.907)
odds ratios (95% confidence interval) ** p-value < 0.001 *p-value < 0.05			

**Table 3.4 (continued)**

	Model 1	Model 2	Model 3
<b><u>Hispanic</u></b>			
<b>female</b>	0.818** (0.527 - 1.268)	0.973 (0.624 - 1.515)	1.066 (0.681 – 1.670)
<b>Teen parent</b>		0.182** (0.066 - 0.503)	0.207* (0.075 – 0.575)
<b>Fam. Poverty</b>		0.531* (0.336 - 0.840)	0.996 (0.559 – 1.654)
<b>Mother's Edu.</b>			1.101* (1.036 – 1.171)
<b>Sibling Edu.</b>			1.272** (1.157 – 1.398)
<b>Siblings</b>			0.971 (0.886 – 1.064)
odds ratios (95% confidence interval) ** p-value < 0.001 *p-value < 0.05			

2015), which means that blacks as a group have had a much longer experience with teenage parenthood, so among blacks we might expect better adaptation to early parenthood and thus a less negative effect of teen parenthood on various outcomes compared to other racial-ethnic groups. This reasoning is modestly supported by the black-white comparison. And among individuals in the NLSY who came of age prior to the time period (mid-1990s) when the teen birthrate among blacks first dipped below that of Hispanics (see Martin et al. 2015), we should expect a similar thing, but this is not what the data suggested with regard to the black-Hispanic comparison. Model 3 in Table 3.4 indicated that the effect of teen parenthood was less negative (but still substantially negative) for Hispanics than it was for whites and blacks. As with the

results about odds of college completion presented previously, these findings do suggest a substantial disadvantage among racial-ethnic minorities, but still it is unclear whether this is due to minority status per se or if other relationships in the data might help explain or account for some of the association between minority status and college completion.

In thinking about how to theorize the negative influence of teenage parenthood on odds of completion, I became curious about whether racial differences in education might be attributable to other facets of people's early lives, such as the unequal likelihood of teenage parenthood across racial-ethnic groups. [As young baby boomers moved into adolescence and early adulthood, Blacks especially were much more likely to become teenage parents compared to whites and Hispanics (Furstenberg 1987)]. Thus, the next subsection discusses some plausible indirect effects in this relationship, and then I assess empirically whether there is a mediating effect of teenage parenthood in the association between minority status and completion. Specifically, I address this question of whether teenage parenthood matters for racial differences in completion by using decomposition models, which are designed to detect such potential effects by comparing coefficients across full and reduced nonlinear (logit) models.

#### *Racial-Ethnic Differences in College Completion: Indirect Effects and a Mediating Effect*

If there are indirect effects at play, what are some straightforward ways to think about them? Theoretically what should we expect regarding potential correlated effects of, say, family poverty, minority status, and college completion? Or what should we expect of potential correlated effects of mother's education, minority status, and completion? While this subsection only briefly discusses some potential indirect effects among the variables defined previously, shortly I will ask and address an empirical question, namely: How much of the negative

relationship between minority status and completion is attributable to or mediated through the influence of other factors? Specifically, is there a mediating influence of teenage parenthood (or of family poverty or parental education) in the minority-college relationship? I pose this as a key question in the context of results presented previously.

Here however I discuss some reasons for suspecting indirect effects, and then briefly theorize them. Next, I posit teenage parenthood and other factors as mediators, and then present results from a decomposition technique developed for assessing mediating effects in nonlinear models. The findings will indicate a strong mediating influence of teen parenthood and (unexpectedly) of maternal education on the association between racial-ethnic minority status and college completion. Finally, I will conclude with some discussion of this apparent empirical connection and of its implications vis a vis existing stratification scholarship.

### *Brief Discussion of Plausible Indirect Effects*

There are reasons to have diverse expectations about confounds or indirect effects in terms of the issue at hand, namely conditional odds of college completion. Previously we saw that controlling for economic disadvantage in individuals' families of origin (most pronounced among blacks) and for parental and sibling educational attainment (most pronounced among Hispanics) seems to attenuate what first appeared to be a substantial negative effect of minority status on odds of completion. Detecting group-level differences of coefficients in separate logistic models, as previously discussed, is not the same thing as detecting correlated effects, mediation, or other types of indirect effects, but it does *suggest* that family poverty and parental/sibling education are at least correlated with minority status and with odds of college completion. One possibility, which aligns with some influential scholarship regarding the relative effects of race-ethnicity and

class background on people's educational fates (e.g., Conley 2001), is that experiencing poverty during childhood could be associated with what merely appears to be a negative effect of racial-ethnic minority status per se on college completion. But an individual's minority status cannot be causally prior to poverty in her family of origin, so as I mentioned this is best thought of as a potential correlated effect, and a relatively straightforward one at that.

There also is reason to suspect that, with eldest sibling's education, which in results discussed previously seemed to have a positive influence on completion, there might also be a correlated effect at play here. Elder siblings can exert a strong socializing influence on younger siblings, inculcate norms or expectations regarding educational achievement, and impart various advantaging traits such as study habits, motivation, or certain kinds of academic aspirations and human capital that get rewarded in formal schooling. Given racial minorities' overall lower levels of educational attainment compared to whites, their eldest siblings' lower levels of education may be detrimental to their educational attainment in any one or more of the aforementioned ways. A similar reasoning also may pertain to mother's education; while we could posit mother's education as a mediator—an individual's minority status can be “causally prior” to her mother's (or her eldest siblings') education if we assume race-ethnicity is exogenous to both respondent and her mother—for illustration purposes here it is more straightforward to presume a correlated effect in the minority-college connection. Finally, considering the negative influence of sibship size on odds of completion, and since blacks and Hispanics tend to have larger families, we might expect that a correlated effect also is at play in the minority-college connection.

Theoretical discussion, as in the foregoing, can help us think through the possibilities, but what can empirical analysis reveal about indirect effects in a model of college completion?

Mediation, which is logically and empirically distinct from confounding, moderation, and covariation, is of particular interest for the analysis presented below. To briefly illustrate the distinction: Mother's education (or number of siblings, etc.) may be correlated with an individual's minority status and with their odds of college completion, so that information about mother's education could improve prediction of completion by race-ethnicity. That is, minority status correlates negatively with mother's education, but in this case mother's education will not substantially alter the nature of minority status's effect on completion when mother's education is included in a model. This is an example of a covariate, but we could pose mother's education as a mediator if we assume that race-ethnicity is exogenous to both respondent and her mother, i.e., if we assume that mother's education is in a plausible causal sequence between predictor and outcome (MacKinnon et al. 2007). (We could posit that a person's race-ethnicity is causally prior to her likelihood of becoming a teenage parent—certain conditions that tend to characterize minorities' lives could increase their likelihood of becoming teenage parents. This fits with two key aspects of causal order, namely temporal and logical precedence; by contrast, we could not posit her race-ethnicity as causally prior to ascriptive conditions such as her mother's education or her family's poverty status.) As will be shown below, diagrams of mediation assume several causal pathways, but two are of particular interest, namely the direct effect of X on Y and a mediated effect by which X indirectly affects Y through a mediator M. The next subsection discusses mediation specifically with an eye to teenage parenthood's potential intervening role in the minority-college relationship.

### *Evidence of Mediating Effects*

Decades of research has shown that pregnancy and parenthood during adolescence are negatively associated with various developmental outcomes for both parents and children, including their eventual educational attainments (Jencks et al 1979; McLanahan and Sandefur 1994; Shaff et al. 2008), though arguments about teenage parenthood's negative effects are subject to a vast body of scholarship and vigorous debate too extensive to summarize here.<sup>39</sup> As displayed previously in Table 3.3, the influence of teenage parenthood among men and women went in theoretically expected directions: For men, becoming a teen parent was associated with substantially lower odds of college completion compared to men who were not teen parents, and among women it was associated with still lower odds (0.157 for women vs. 0.178 for men). At all ages, women tend to bear more responsibilities for childrearing than men do in the United States, so it was somewhat surprising that teen parenthood seemed to have only a slightly more suppressive effect on women's chances of completing college over time. (Another brief reminder about how I describe results: While I use the terms "odds of completion", "odds ratio," and "chances of completion," these terms are potentially misleading and should be read as shorthand for the discrete-time outcome. What I refer to as the "odds of college completion" actually reflects a comparison of the relative chances for completion at each age across individuals' lives.)

By contrast, the connection between racial-ethnic minority status and odds of college completion seems less straightforward. The apparent negative influence of minority status on people's odds of completion may be "real" in that there could exist a set of cultural practices, family values and attitudes toward education, shared ethnic experience, or other unobserved

---

<sup>39</sup> Those who are better educated tend to marry later and have children later than those with less education (Blossfeld 1995; Liefbroer and Corijn 1999), but the relationship between early family formation and educational attainment may operate in more than one direction: Early family formation may negatively affect educational attainment, but it is also plausible that unsavory school experiences (or any number of contextual factors) during adolescence may positively influence adolescent family formation and childbearing.



characteristics unique to racial-ethnic minorities that reduce their odds of completion relative to whites. However, it could be that the negative influence of minority status on completion is attributable to how the values of other variables are patterned, or that it operates through other variables' effects on completion. Specifically, given the models discussed previously, it could be that the overall negative effect of racial-ethnic minority status on college completion operates indirectly by its positive influence on a person's likelihood of becoming a teenage parent. The assumption here is that minority status can be considered causally influential on teenage parenthood, and one can theorize it in various ways. For example, minority adolescents, with their greater odds of living in socioeconomically disadvantage families and neighborhoods, may perceive less of an incentive to "wait until after college" or "wait until marriage" to have children, plausibly due to lower aspirations for college or the low availability of marriageable men in black American communities. It is also plausible, and strongly suggested by a long tradition of research (e.g., Furstenberg 1987; Martin et al. 2015), that black adolescents are more likely to perceive a relatively low pay-off, so to speak, in delaying childbearing as compared to other racial groups for whom there is a generally brighter socioeconomic future (for whom local labor markets are more robust, or who have optimistic college aspirations, or who live in neighborhoods of higher socioeconomic status, and so forth).

With the same sample of individuals examined in this section's previous discussions, I attempted to detect mediation effects<sup>40</sup> by comparing coefficients across full and reduced nonlinear (logit) models. Using the *-knb-* program for logit models, first I decomposed the total effect of racial-ethnic status on risk of college completion into its direct component and its

---

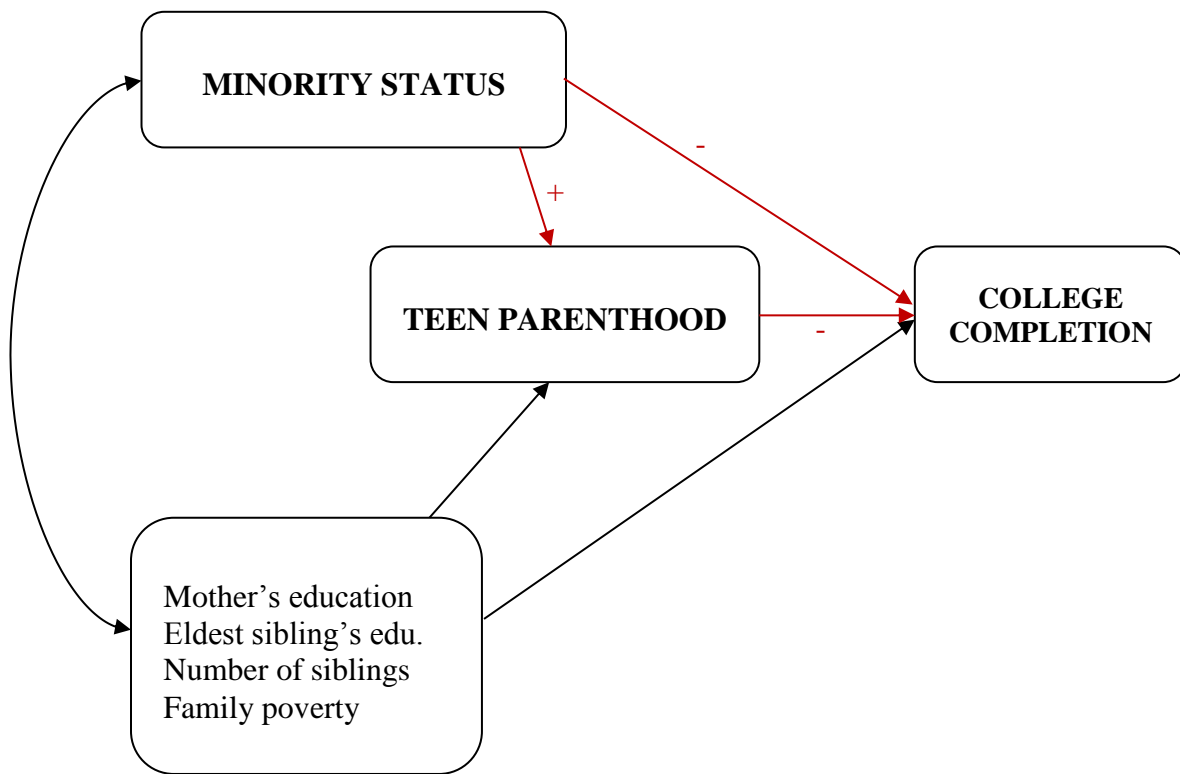
<sup>40</sup> In contrast to moderator variables which are hypothesized to affect the direction or strength of association between predictor and outcome (as in statistical interaction), mediator variables can help to explain or account for the strength or direction of a relationship between predictor and outcome.

indirect component running through teenage parenthood.<sup>41</sup> Then I ran this same procedure and added family poverty, mother's education, number of siblings, eldest sibling's education, gender, and age as controls.

Figure 3.2 presents a visualization of the hypothesized relation between racial-ethnic minority status and college completion, with arrows representing causal directions among the independent, dependent, and mediator variables, and positive/negative signs indicating directions of influence. For illustration purposes, the three red arrows are meant to indicate the proposed mediation in this example, and the curved double-ended arrow simply indicates that minority status is seen to be correlated with the corresponding variables in the other box. To start off using this diagram for reference, I would expect individuals who became teenage parents to have lower odds of college completion at any age, and for reasons described above this relationship could account for some or much of minority status's negative association with completion. In terms of mediation, the diagram posits that minority status negatively affects completion and positively relates to teenage parenthood, which itself also negatively impacts completion. In this relatively example, at least some of the apparent negative influence of minority status *per se* on college completion is presumed to run through (or be partially explained by) teen parenthood's "more negative effect" on completion among blacks and Hispanics compared to whites.

---

<sup>41</sup> The magnitude of logit coefficients is difficult to sensibly interpret, so odds-ratios were calculated to ease interpretation.

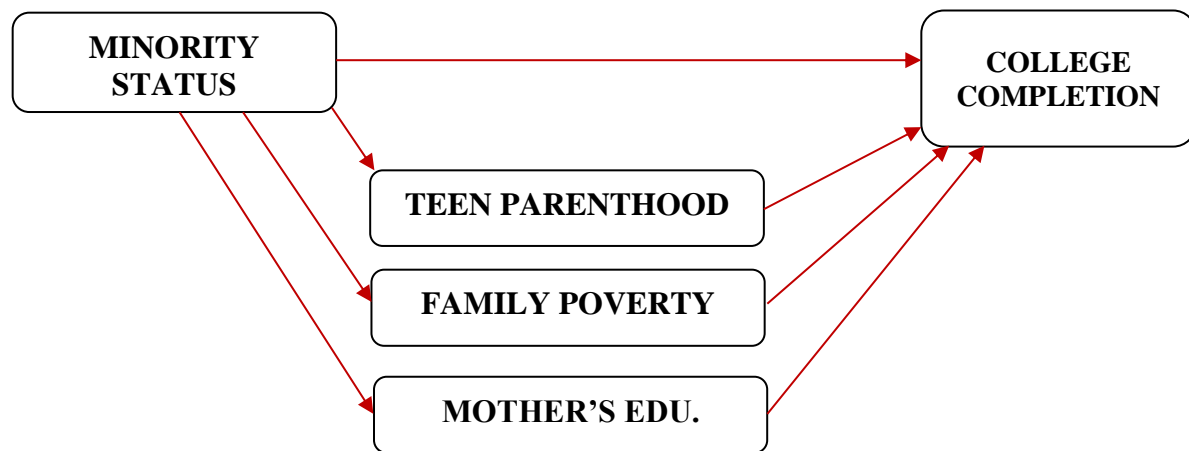


**FIGURE 3.2**  
**Single-mediator model: Racial-ethnic minority status, family background, and conditional odds of college completion (with teenage parenthood as proposed mediator)**

Additionally, and as will be explored with the decomposition results presented below, one can imagine a number of plausible relationships, including pathways that pose other or additional mediator variables (teen parenthood, family poverty, mother's education, etc.). In short, a multiple mediation model allows for determining the relative magnitudes of the indirect effects associated with all mediators. A multiple mediator model also is advantageous because it provides a better way to assess the plausibility of competing theories within one model than could be done by simply running several single mediator models separately.<sup>42</sup> With reference to Figure 3.3, the specific indirect effect of minority status on college completion through teen

<sup>42</sup> As has been discussed in the methodological literature, running several single mediator models separately could lead to biased parameter estimates (Preacher and Hayes 2008).

parenthood represents teen parenthood's mediating effect conditional on the inclusion of other mediators in the model. (Since any given individual's racial-ethnic status cannot be causally influential on mother's education or poverty in family-of-origin, in this case I conceptualize racial-ethnic minority status as exogenous to both the individual and to her mother/family.)



**FIGURE 3.3**  
**Multiple mediator model: racial-ethnic minority status and conditional odds of college completion with three hypothesized mediators**

Referring back to the visualization presented in Figure 3.2, minority status can be considered causally influential on teenage parenthood and on college completion, and this can be seen empirically. In single-mediator models not shown here, the total effect of Hispanic ethnicity on conditional odds of completion was only slightly larger in magnitude than its direct effect (1.068 vs. 1.066), and only 2 percent of the Hispanic effect on completion ran through teenage parenthood. That is, teen parenthood did not seem to account for much of the relationship between Hispanic ethnicity and college completion, at least not in that initial modeling scenario. This finding starkly contrasted with teen parenthood's mediating role among blacks, for whom slightly more than half of the total effect of minority status on college completion was accounted for by teen parenthood's negative effect on completion. This suggested a substantial intervening

influence for teen parenthood among blacks but not Hispanics, and it suggested a mediating effect for family poverty among both blacks and Hispanics.

If the overall negative minority-college relation seemed to operate indirectly by minority status's positive influence on black individuals' greater likelihood of teenage parenthood and of Hispanics' and blacks' greater likelihood of childhood poverty, then what happens when other possible mediators are added to the model? After including various combinations of mediators, it became apparent that teenage parenthood, family poverty, and mother's education (but not eldest sibling's education or number of siblings) were the three factors that seemed to play the most powerful mediating role as a group. Table 3.5 displays estimates for the decomposed effects (total, direct, and indirect) of racial-ethnic minority status on individuals' conditional odds of completing college as it is proposed to run through these three mediators. (White is the reference category for race-ethnicity, and controls are indicated at bottom of the table.) First, what is labeled "overall mediation percentage" indicates how much of the total effect of minority race-ethnicity was due to the confounding of the three mediators combined. For blacks and Hispanics, a substantial share (73 percent and 91 percent, respectively) of the overall association between minority race-ethnicity and conditional odds of college completion was accounted for by including teen parenthood, family poverty, and mother's educational attainment as mediators.

Specifically which mediators contributed most to the confounding between minority status and college completion? An indicator for the relative mediating influence of teen parenthood, family poverty, and mother's education is displayed in the "Mediation Percentage"

**Table 3.5**  
**Multiple-Mediator Model: Decomposed Effects of Minority Status on Conditional Odds of College Completion (with three mediators)**

N = 97,053 person-years

	Log Odds	Odds Ratio (95% C.I.)
<b><u>Black</u></b>		
Total effect	-0.316	0.729 (0.600-0.885)
Direct effect	-0.086	0.918 (0.757-1.113)
Indirect effect	-0.230	0.794 (0.749-0.842)
<u>Overall Mediation %</u> 72.93		
<b>Mediators</b>		<u>Mediation Percentage</u>
Teenage Parenthood	-0.172	54.56
Family Poverty	-0.061	19.29
Mother's Edu.	0.003	0.92

**Table 3.5 (continued)**

<b><u>Hispanic</u></b>		
Total effect	-0.821	0.440 (0.342-0.566)
Direct effect	-0.069	0.934 (0.730-1.195)
Indirect effect	-0.752	0.471 (0.427-0.521)
<u>Overall Mediation %</u> 91.64		
<b>Mediators</b>		<u>Mediation Percentage</u>
Teenage Parenthood	-0.077	9.41
Family Poverty	-0.044	5.34
Mother's Edu.	-0.631	76.89

Controlling for covariates: gender, age, number of siblings, family income in 1979, eldest sibling's education, and "intact" family structure during adolescence.

column. This shows how much of the total minority-college effect is due to confounding of each mediator, conditional on the inclusion of the other mediators and controls. Here we see several notable results, the first of which is that, among blacks, teenage parenthood and family poverty together accounted for nearly three quarters (54 and 19 percent, respectively) of the overall confounding, and mother's education did not seem to matter. The mediation percentage for teen parenthood among blacks was nearly six times greater than its share of mediation among Hispanics (54.56 vs 9.41 percent), and experiencing poverty during childhood was four times more influential as a mediator for blacks compared to Hispanics.

As we saw above, mother's education intervened very little for blacks but was a much more powerful mediator for Hispanic ethnicity, accounting for more than three quarters (76.89 percent) of the Hispanic-college association. (This will be discussed in greater detail below.) Also among Hispanics, teen parenthood accounted for nearly twice as much (9.41 vs. 5.34 percent) of the confounding as family poverty did. In sum, although the three mediators together accounted for around three quarters of the overall black/Hispanic effects on completion, the magnitude of mediation was much larger for teen parenthood and mother's education compared to family poverty's intervening effect. And even though teenage parenthood seemed to matter somewhat for Hispanics, it mattered much more for blacks.

As mentioned previously, much research has shown that black adolescents' greater likelihood of socioeconomic disadvantage may lead them to perceive less incentive to "wait" on childbearing, perhaps due to lower aspirations for college, lower educational expectations from parents, the proximal influence of low-achieving peers, or (for women) the relatively low availability of marriageable men in black American communities. Black adolescents may also perceive a relatively low pay-off in the labor market for delaying childbearing. These

possibilities were raised in my previous theoretical discussion, and they suggest at least a plausible interpretation of why so much (more than half) of the supposed effect of being black on college completion ran through teen childbearing's unfavorable influence on completion.

By contrast, teenage parenthood gets credit for only a small percentage of the overall effect of being Hispanic on people's odds of college completion, and this paled in comparison to how much mother's education intervened. Strikingly, mother's education accounted for more than three-fourths (76.89 percent) of the Hispanic-college connection. This magnitude of mediation, especially the stark racial-ethnic difference in its importance, was surprising and is not something about which I initially theorized or hypothesized, so the result requires some post-hoc framing and interpretation. Sociological theory and existing research can help here, especially in terms of the potential causal connection between parental aspirations and educational achievement. Much existing scholarship supports the assumption that children from Hispanic families experience different kinds of parental socialization and different social psychological inducements to succeed in school, and that better-educated mothers would have a stronger effect on their children's educational aspirations compared to their black counterparts. There is a fairly well documented link between parents' stated educational expectations and how well children do in school, and this connection seems to vary substantially by race-ethnicity, with Hispanic parents reporting much higher educational aspirations for their children than whites or blacks (Behnke et al. 2004). There also is a body of research showing that Hispanic parents tend to believe that schooling is a primary pathway for their children's upward social and economic mobility, they believe this to a substantially higher degree than black parents do, and even Hispanic parents who have very *low* educational attainments tend to report very *high* educational aspirations for their children compared to black parents (Goldenberg et al. 2001). So no matter



the level of a respondent's mother's education, some cultural dimension of being raised in a Hispanic family could be influencing their likelihood of completing college. A related line of research has suggested a social learning effect—according to several studies, Hispanic parents not only have more influence on their children's educational decisions than any other racial-ethnic group but, relative to blacks and whites, children from Hispanic families have higher academic motivation if their parents had a strong ability to help with schoolwork (Qian and Blair 1999; Plunkett and Bamaca-Gomez 2003; Behnke et al. 2004). We can reasonably assume that mothers with more years of education would have higher educational aspirations for their children and that they would be better able to help with school-related tasks, so it is possible that Hispanic individuals whose mothers had stronger academic skills—as indicated by more years of education—would themselves have better odds of completing college due (in part, at least) to this indirect effect. In short, while I cannot make empirical claims about individuals' interactions with their mothers during the years when they lived at home, or even about what their mothers' aspirations were for them during adolescence, at the very least we have a theoretical framework to draw on in making some sense of a rather surprising result.

### **Concluding Discussion and Ideas for Future Work**

Here I briefly summarize this section's main findings, offer some comments on what was missing from the study's theoretical considerations and empirical analyses, and then conclude by mentioning possible directions for future research.

An initial descriptive finding was the somewhat unexpected heterogeneity in timing of college completion; a basic result, yet noteworthy in terms of one conventional life course theoretical assumption (not much variability in age at which people complete college) because it

suggests that post-secondary education among members of this baby boom cohort may not be as distinctly age-differentiated as some might assume. Put another way, among members of an earlier generation for which one might assume “traditional” ages at college completion, we should not consider the *early twenties* to be the period that is exclusively associated with it. Even among a sample of adults who came of age several decades ago when there was presumably not much variability in the ages at which people completed college as compared to more recent generations, the individuals examined here were completing college during periods of life that surely overlapped with other important events and transitions such as marriage, childrearing, and so forth.

When I examined conditional odds of college completion using an event history framework and logistic regression, I found (unsurprisingly) that the experience of family poverty during childhood had a negative influence on people’s odds of completing college at any age, and that the inclusion of this measure in the models attenuated the negative impact of racial-ethnic minority status. (In that part of the discussion I suggested that the family poverty effect may operate indirectly through other features of people’s early lives.) In considering the possible influence of siblings, initially I expected that the presence of more siblings in their families-or-origin would reduce individuals’ odds of completing; additional siblings can dilute resources like time, money, or energy that are available to each child. Somewhat surprisingly, results indicated only a slight negative effect of number of siblings on completion. In examining the presence of siblings, my findings aligned with the general received wisdom among stratification scholars (that sibling’s education should favorably influence people’s chances of completing college), but not in nearly as pronounced a way as theoretically expected. And while eldest sibling’s education *per se* did not seem very influential, its inclusion in the models was associated with a change in

the effect that racial-ethnic minority status had on people's conditional odds of completion. With regard to racial-ethnic minorities, the findings showed an unsurprising disadvantage among them and among individuals whose families experienced poverty during their childhoods. In short, few surprises there.

To inquire more into these relationships, I discussed some expectations about indirect effects and used mediator decomposition models to help clarify the apparent negative association between minority status and college completion. I suspected that teenage parenthood and family socioeconomic disadvantage could have the most significant intervening effect in the minority-college relationship, and the evidence suggested strong mediation effects of teen parenthood and family poverty on the black-college association (but, as I discuss below, there was a different story for the Hispanic-college connection). I then invoked some basic tenets of theory from the stratification tradition so as to make sense of this apparent mediation.

One of the more striking findings was that, while teenage parenthood seemed to be a strong mediator of the black-college relation, this was not the case for Hispanics in the sample. Bluntly put, teen parenthood paled in comparison to how much mother's education intervened in the Hispanic-college association. Initially I had no theoretical reason to expect that this would be the case, so I sought clarity in existing scholarly claims about racial-ethnic differences in family life. In the U.S., Hispanic parents seem to hold stronger beliefs about the potential for formal education to transform their children's lives. Hispanic parents tend to believe that schooling is a primary pathway for their children's upward social and economic mobility and they believe this much more strongly than black parents do. I posited that various cultural dimensions of being raised in a Hispanic family may affect the likelihood of children's college completion in a way that it does not for blacks. There is also the possibility of a social learning effect, at least insofar

as Hispanic parents have a stronger influence on their children's educational decisions than any other racial-ethnic group, and insofar as children in Hispanic families have higher academic motivation when their parents have a strong ability to help with schoolwork. The assumption here is that mothers with more years of education would have higher educational aspirations for their children and that they would be better able to help with school-related tasks. In short, I argued for the possibility that Hispanic individuals whose mothers had more years of education would themselves have better odds of completing college due (in part, at least) to this indirect effect. I cannot make empirical claims about individuals' early family lives or parental interactions, nor can I know what their mothers' aspirations were for them during adolescence, but at the very least it was worth drawing on theory and existing scholarship to try and make sense of an otherwise surprising result from the decomposition models.

It is impossible for a single study to cover all the proverbial bases, but since this study is broadly about educational stratification (a diverse literature in itself), its limitations deserve at least some mention. One of those limitations has to do with the huge number of possible predictors for educational success. At the individual level, academic preparation is an important determinant of college completion; several studies have identified the types of high school coursework students undertake as an important predictor of their ability to do well in college. Though Adelman (1999; 2006) used a longitudinal dataset other than the NLSY and did not inquire into the timing of college completion, his important studies of a cohort who were tenth-graders in 1980 found that their academic backgrounds—defined by detailed measures of academic content, types of courses taken, high school curriculum intensity, class rank, and GPA—were crucial factors in college success. There also is evidence that the proportion of high schoolers who are “prepared” for college is much smaller among blacks and Hispanics than

whites (Greene and Foster 2003), and indeed research has shown substantial gaps in “college readiness” by racial-ethnic status and class background, at least as it is indicated in standardized test scores (Jencks and Phillips, 1988). One limitation of my work here is that much of the detailed information one would need in order to examine the aforementioned factors is not available in the NLSY 1979 surveys.

Another limitation is that my analyses included only a few coarse measures for family-level and parental characteristics, and it would have been beneficial if I had been able to include other attributes such as parenting style, individual psychological disposition, adolescent criminal behavior, and adolescent social networks, as these also have been identified in the literature as influential of children’s life chances including their likelihoods of completing college (Mayer et al. 2004). Another interesting angle, indicated by parental income’s correlation with parental and children’s cognitive skill levels, is that at least part of the gap in risk of college completion between children from higher income families and those from lower income families is accounted for by differences in parents’ and children’s cognitive skills (Mayer 2010, 23). As a more general matter, much research has shown that knowledge of children’s early lives is crucial to understanding the various effects of social inheritance (Esping-Andersen 2004; Erikson & Goldthorpe, 1992; Shavit & Blossfeld, 1993). With all this in mind, one approach that would add substantively and explanatorily would be to examine NLSY 1979 respondents and control for individuals’ cognitive test scores during adolescence. This could be done using the Peabody Individual Achievement Test (PIAT) and the Armed Services Vocational Aptitude Battery (ASVAB), both of which are available in this particular dataset. Again, this would not measure the more detailed aspects of social networks or parenting style previously mentioned, but it would be an improvement at least.

Other studies (e.g., Arum and Beattie 1999) have suggested that, in attempting to model college completion, we should look to other factors such as the nature of people's intimate social bonds, the quality of their peer relationships, and participation in conventional activities that are associated with a "college readiness" mentality. This line of research also has identified high school characteristics (school climate, availability of vocational/occupational coursework vs. college-prep curricula) and student/teacher ratio (smaller ratio has been linked to higher overall educational attainment) as key factors in college success. Still others have pointed out the importance of contextual variables such as the robustness of county governments, area median income, and neighborhood poverty (e.g., Lobao et al. 2012) in explaining later-life outcomes such as educational success.

In terms specifically relevant for life course theory and research, an important consideration for future work using NLSY would be to attempt to trace cumulative disadvantage beginning in people's early lives (mentioned previously in a footnote to the current work). Inequalities that begin earlier in people's lives—such as childhood illness, poor quality kindergartens, or school tracking in later grades—can constrain people's chances of acquiring what has been referred to as "life course capital," or the overall stock of economic, social, and personal resources that people draw upon in dealing with uncertainties, expected life transitions, and unexpected events (Pallas 1993 and 2004; O'Rand 2001). Also deserving of closer scrutiny than was attempted here, and hinted at previously, are the various dimensions of family socioeconomic conditions during childhood (family poverty as was examined here, but also low parental occupational status, low parental wealth, etc.), many of which seem to have long term negative effects on people's cognitive and psycho-social development (McLeod and Almazan 2003), socioeconomic attainment (Elman and O'Rand 2004), health in mid-adulthood (O'Rand

and Hamil-Luker 2005), and even on the odds of mortality during later adulthood (Hayward and Gorman 2004). Any attempt to analyze the effects of parental wealth would need to use data other than the NLSY79 (as it does not provide information for respondents' parental wealth).

Much research in the life course tradition attempts to understand the potential links between childhood conditions and subsequent life course trajectories, primarily in educational, occupational, and familial domains, and a core tenet of life course theory is that advantages during childhood tend to translate into resources—psychological, social, and economic—that then lead to further relative advantages. Educational attainment clearly is one such advantage. As we saw in this study, this advantage seems to derive, at least in part, from other features of people's early lives (as with family poverty), or of their entire lives (as with race-ethnicity), or through an indirect influence of these factors (as with teenage parenthood and race-ethnicity). Educational stratification, including the relatively straightforward phenomenon of differences in people's chances at completing college that was examined here, obviously can be shaped by individual idiosyncracies. But this work confirms a common maxim in the literature, specifically that stratification also is driven by patterns of advantage and disadvantage starting early in the life course.

Finally, one particularly interesting consideration for future work with the NLSY is that of racial-ethnic and social class variation in the *timing of the transition from high school to college*, and whether delayed college enrollment lowers people's chances of timely college completion. At first this may sound non-sensical, for surely delayed enrollement necessarily delays eventual completion, but as the normative age at college completion shifts upward—today nearly half of college students enter and complete their degrees at “non-traditional ages”—we should ask about the effects of delay on the odds of people completing by, say, age 24 or 25

rather than what were for so long the normative ages (22 and 23). An emerging body of research is beginning to show that when people transition from high school to college may be a key predictor of whether they complete a four-year degree (Adelman 2006; Bozick and DeLuca 2005; Goldrick-Rab and Han 2011). However, as with much other scholarship in the field of educational stratification, existing research has not adequately asked about the “when” of college completion even though it examines the when of college *enrollment*, nor have scholars investigated racial-ethnic differences in delayed enrollment or its influence on eventual completion. Indeed, to my knowledge no research has examined how the timing of enrollment may affect the timeliness of people’s college completions. Social processes like this (of the general form: *how timing of an early educational transition affects timing of future educational transitions*) have been well theorized but not adequately examined empirically in life course scholarship, so studying this would be timely considering that recent American generations are increasingly undertaking non-traditional educational trajectories with norms shifting toward completion into the middle and later 20s. (Ideally the NLSY’s 1997 dataset—a completely different sample than the 1979 data examined here—would be better suited to such analysis.) Whether people who postpone enrolling in college one year (or two years, or three) after finishing high school are more or less “on time” in completing college prior to, say, age 26 or 27 is an intriguing empirical question, the study of which would further our understanding of educational outcomes.



## SECTION 4

### **“BETTER TO DO (SOME) THINGS EARLY”: TIMELY 4-YEAR COLLEGE COMPLETION, TEEN PARENTHOOD, ADOLESCENT EMPLOYMENT, AND WEALTH MOBILITY IN MID-ADULTHOOD**

This dissertation’s two previous studies conceptualized the timing of educational completion as an outcome, but my final study will theorize and analyze it (along with other factors) as consequential for people’s socioeconomic fates. Researchers routinely study educational completion as an outcome or predictor, but less commonly have they thought of “timely” completion (earning a college degree by, say, age 24) as a predictor for later life conditions such as socioeconomic mobility over people’s adulthoods. Furthermore, sociologists in general but more particularly users of the NLSY have given relatively little attention to wealth compared to income or occupational position as a dimension of socioeconomic standing. Thus, one of the current study’s contributions is to study wealth mobility in the context of timely educational completion and other early-life transitions. With the data available in NLSY, I am able to examine wealth mobility over an important period in people’s mid-adulthoods (from age 24 to age 44). In doing so, I add to an emerging literature in stratification and mobility studies by examining how timely college completion and other important early life transitions, such as adolescent employment and teenage parenthood, may influence people’s odds of experiencing wealth persistence (staying within upper or lower quartiles of the wealth distribution) and wealth mobility (into or out of the upper/lower quartiles) by the time respondents reached their mid-40s.

A vast literature continues to document the effects of socioeconomic status during childhood on various measures of achievement (e.g., Conley 1999; Yeung et al 2002; Bradley and Corwyn 2002; Torche 2011). As shown previously in Section 3 of this dissertation, aspects

of family background such as income poverty during childhood and other factors such as racial-ethnic differences in teenage childbearing matter for the timing of individuals' of college completion. Those findings aligned with other studies that point to the disadvantages of poverty and minority status for a range of educational outcomes (e.g., Roksa and Velez 2012). However, another aspect of socioeconomic status, namely wealth—the distribution of which is much more unequal than income—has received relatively little attention in this particular body of literature. For decades, studies of socioeconomic mobility focused mostly on educational attainment, occupational prestige, income, and cognitive skills, with negligible emphasis on wealth until Oliver and Shapiro's (1995) and Conley's (1999 and 2001) important studies demonstrated its effects on various dimensions of stratification.<sup>43</sup>

Wealth inequality now receives considerable scholarly attention, especially in studies that focus on the intergenerational transmission of socioeconomic advantage, but only recently have scholars suggested that certain traits and experiences that characterize people's early lives, such as adolescent employment and timely educational completion, may help to explain financial inequalities among them later in adulthood (e.g., Keister 2005; Painter 2010). While inter- and intra-generational mobility are both of interest, I focus here, for reasons mentioned below, on *intra-generational* mobility.

I mentioned previously that the focus of this final section represents a substantial shift compared to Sections 2 and 3, in which I analyzed people's chances of completing college at any given age. Here instead I explore some basic yet under-examined questions about wealth among individuals in the NLSY 1979 sample. The main outcomes of interest here are wealth levels and intra-generational wealth mobility (shifts in people's positions within the wealth distribution

---

<sup>43</sup> As with most subfields in stratification scholarship, the status attainment literature is vast. Some early exemplars of the tradition just mentioned are Jencks (1972); Bielby et al. (1977); Hauser and Featherman (1976); Hauser and Mossel (1985); Grusky and Hauser (1984); Hauser and Sewell (1986); and Hout (1988).

over the life course) within respondents' lifetimes. Specifically the analyses begin by documenting wealth levels and wealth inequalities over the time when individuals' aged from emerging adulthood (24 years old) into mid-adulthood and the beginning of their peak earning years (44 years old). This particular cohort was chosen out of the need to capture individuals after they left their parents' households but prior to when they retired (retirees tend to spend down their wealth or transfer it *inter vivos*, and they are more likely than pre-retirees to have inherited at least some wealth; either of these scenarios could distort the measurement of accumulated wealth and therefore wealth mobility).

After presenting descriptive wealth statistics for the sample, I identify some racial-ethnic patterns in wealth levels, wealth mobility, and wealth volatility as individuals aged. This study then connects substantively, not methodologically, to the general emphasis in Sections 2 and 3 on educational completion, but here a key question is whether individuals who completed their four-year college degrees "on time" (prior to age 24) were more likely to experience upward wealth mobility by mid-adulthood than those who completed later. In addition to modeling the influence of timely college completion on wealth mobility within an individual's lifetime, I address the life course theoretical tradition more directly than was done in previous sections by exploring one of that tradition's core themes known as "transitions to adulthood." Employment during adolescence and teenage parenthood are two of the most common life course transitions (marriage is the other transition most often studied in life course research). Specifically I examine the degree to which teenage parenthood and adolescent employment, respectively, influence people's wealth mobility by mid-adulthood.

Before presenting the analyses, first I offer some context by summarizing contemporary scholarly discussions of wealth, wealth mobility/volatility, and patterns of adolescent

employment in the United States. The summary discussion is meant to justify my scholarly concerns and to foreshadow the research questions listed subsequently. After laying out the research questions, I describe this study's data, methods, and measures. The ensuing results and discussion subsections should contribute to the stratification literature in three ways, the most basic of which is that I describe some basic patterns of wealth using many years' worth of underused data from an important longitudinal dataset (NSLY). I add to existing social science knowledge by addressing descriptive questions about wealth levels and wealth mobility among young baby boomers between the mid-1980s and the mid-2000s. This study also contributes to ongoing debates about the connection between certain aspects of status attainment and socioeconomic mobility, specifically by modeling the relations among adolescent employment, timely college completion, and wealth mobility over a substantial portion of the life course.

### **Focus on Wealth**

Research in the status attainment and human capital paradigms has focused largely on family income dynamics, and most economic and sociological models conceptualize income (a more inclusive variable than earnings) as the "sum of returns to the factors of production a person brings to the market, such as cognitive functioning and education" (Bowles, Gintis, and Groves 2005, 4). But the returns to a person's "factors of production" encompass much more than a person's income or earnings. The focal concern here, wealth, has characteristics that arguably make it a meaningful component of the returns to the factors of production that someone brings to the labor market, especially over the long term. This is best reflected by the important differences between income and wealth as sociological phenomena. Unless income from employment gets invested, secured as cash savings, or otherwise transformed into some type of

wealth, it cannot be directly transferred across generations (except indirectly, perhaps, by enhancing the cultural and human capital of the next generation). When derived from employment, income ends when its recipient becomes unemployed, retires, or dies. Obviously individuals do not sell their labor power after death, but wealth—accumulated indirectly as a function of the returns to a person’s “factors of production”—can persist and grow in magnitude after a person’s death, thereby affecting the psychological and material lives of many descendants.<sup>44</sup>

Wealth is intriguing in others ways. Its possession and accumulation (or even a low-wealth person’s expectation about receiving an inheritance, or a high-wealth person’s willingness to take financial risks or to invest in their children) can endow individuals and families with important psychological, familial, and material advantages beyond those that could be gained through employment income alone. Wealth can yield investment returns (for various types of assets) or be used as collateral in small business loans, and it can enable families to handle expected or unexpected expenses associated with illness, unemployment, and caretaking of elders (Torche and Costa-Ribeiro 2012). Obviously in some situations income also can be useful in these regards (especially in the rare cases of very high-income families who own little wealth), but for most Americans income depends mostly on employment, thus as mentioned previously it can wither with downward occupational mobility or vanish with prolonged unemployment. Wealth, on the other hand, can originate from many sources, be held and expanded in stock or equity form, be drawn upon (turned into income, in effect) to meet a family’s needs, and it can facilitate various intergenerational advantages through inter-vivos transfers and inheritances.<sup>45</sup>

---

<sup>44</sup> Of course, a family’s wealth can also *decrease* for many reasons such as retirement, medical crisis, inter vivos transfer, sudden unemployment, and so forth.

<sup>45</sup> As with so many social phenomena, various unobserved social factors and individual characteristics may confound the attempt to predict wealth accumulation or mobility during people’s lifetimes. These include inter vivos

Within and across generations, wealth endows a family with at least some protection from unplanned expenses, sudden income shocks, and other adverse life events.

### *Wealth and the “Timeliness” of Educational Completion*

Beyond documenting the basic wealth dynamics in NLSY, this study seeks to account for wealth mobility as people age, and partially it does so with reference to one of the major substantive concerns examined previously in sections 2 and 3, namely the timing of educational completion and specifically the sociological significance of completing college “on time” as compared to doing so “off time.” Some explanation is appropriate here: When people from different socioeconomic backgrounds attend and complete college (at whatever ages) they at least have in common “the college experience,” but substantively their experiences may differ greatly on the path to completing their degrees, with some paths being speedier than others. Research has shown that people from lower socioeconomic backgrounds are more likely to experience “interrupted” college careers than those from higher socioeconomic backgrounds, and such interrupted pathways negatively affect the likelihood of timely college completion (Goldrick-Rab 2006). The analyses presented previously in Section 3 align with other research (Elman and O’Rand 2004), which found that people from higher-income families tend to complete college earlier than people from lower-income families. And as has been well documented, earlier four-year college completion is associated with higher lifetime wages and even higher wages during early adulthood (this effect might be partly spurious; i.e., the behaviors and advantages that led to earlier college completion could also be causally associated with higher wages in adulthood). There are good reasons to expect that those who completed “on time” (prior to age 24) would

---

transfers, the development of risk-averse financial habits or a general “savings orientation,” being socialized into a predilection toward risk-taking behaviors, and so forth (see Bowles et al. 2005 for a comprehensive summary of these complexities).

also experience greater wealth mobility in adulthood into their peak earning years compared to those who were “off time” in completing college. The simplest reason is that people who completed college on time are better positioned to move into higher-wage employment earlier in their lives and to start saving or investing earlier, which should increase their overall wealth accumulation in the long run. At the psychological level, various traits, such as a general savings orientation or a risk-averse financial disposition learned from a parent, may be associated with other behaviors, attitudes, or capabilities (e.g., aspirations, optimism about the future, greater intelligence) that make people more likely to complete college on time and that also are associated with a propensity to begin saving early in life (greater intellectual-cognitive capacity, which is partly inherited genetically and partly induced socially, could matter here too). It also is plausible that, by dint of completing on time with many of their peers, traditional-age completers will be perceived more favorably by potential employers. The status advantage of timely completion itself, regardless of type of degree or institution attended, may boost people’s chances of success in the labor market, which could increase their long-run chances of building wealth and holding onto wealth.

### *Some Key Trends and Disparities in Wealth*

Social scientists (e.g., Painter 2011; Conley 2008) typically conceptualize wealth as a combination of an individual’s or family’s liquid financial assets (stocks, bonds, trusts, retirement saving, cash savings), illiquid holdings such as automobiles or business equity, and residential assets (home value minus value of the mortgage minus property debt). Wealth in the United States is distributed much more unequally than income. It is well documented that wealth “displays greater racial disparities than any other socioeconomic measure [and] these differences

have grown since the civil rights era” (Yeung and Conley 2008, 303). Using various data sets, researchers have shown that racial disparities in family wealth hold even when taking into account income and other demographic characteristics (Oliver and Shapiro 1995; Wolff 1999).

Racial-ethnic inequalities in wealth ownership continue to grow in American society (Wolff 2007; Oliver and Shapiro 2006) but they may be partially accounted for by differences in family structure or family size. Using NLSY data, Keister (2000) examined racial differences in the structure of family of origin and family in adulthood, and the effect these differences had on wealth accumulation. She found that large family size and family disruptions during childhood were negatively associated with wealth accumulation in adulthood (family size was a more important factor in wealth accumulation for whites than for blacks or Hispanics, and family disruption was most strongly related to wealth outcomes for Hispanics). The dilution of financial resources within families (and across extended families, especially among blacks and Hispanics) also seems to influence people’s wealth accumulation in adulthood. Keister’s study also showed that, compared to their white counterparts, even black families who have at least some wealth will tend to spend down whatever wealth they have within a generation—perhaps to assist indebted relatives, to aid extended family members, or to support a small family business—and adult children in black families are less likely to benefit from bequests. Racial-ethnic differences in the receipt of financial inheritances help to partially explain why the average difference in wealth between black and white households is larger than the average difference in income between these two racial groups. Menchik and Jianakoplos (1997) documented the greater likelihood of white households receiving an inheritance than black households. Controlling for other factors which contribute to racial differences in wealth, they estimated that financial inheritance may account for between 10% and 20% of the average difference in black-white



household wealth. Studies like this do aid our understanding of wealth accumulation, but they do not speak directly to the question I address below, namely how to understand racial differences in wealth mobility over the period in people's lives when their earnings tend to peak.

Life course transitions, such as getting married or having children, affect wealth dynamics. Marriage generally improves individuals' wealth standing and wealth stability (e.g., Hao 1996; Wilmoth and Koso 2002), but having children seems to have mixed effects on adults' wealth. Children are expensive, and parents may react in various ways to having children or to the prospect of having children. The arrival of children may force parents to dilute or spend down wealth and increase current consumption. Alternatively, the expectation or arrival of children might motivate more advantaged individuals to save money and acquire assets in anticipation of future expenses associated with having children. Findings on the causal connection between number of children and wealth in later life are not definitive (Painter and Shafer 2011).

There are many ways to theorize the presence of children and their effect on people's wealth, such as how the presence of children seems to affect racial differentials in wealth mobility. This would connect to some of the earlier discussion of racial-ethnic differentials in college completion (except in this case, timely completion is conceptualized as a predictor rather than an outcome). As an example, racial-ethnic minorities start out with lower levels of net worth than whites (e.g., Conley 1999; Oliver and Shapiro 2006), they are more likely to become parents as teenagers, less likely to complete college, and they tend to have more children than white families. Considering the latter pattern especially, we might reasonably expect that the financial resources racial-ethnic minorities possess when they have children will be become more "diluted" or spread out across family members. Thus, one plausible hypothesis is that the

link between number of children and wealth *accumulation* should vary by race-ethnicity. But this study's focus is wealth *mobility*, so beyond the question of wealth accumulation, might number of children have an influence on people's chances of experiencing wealth mobility as they enter their mid-40s? This is just one of many ways to theorize, but the empirical question still remains as to whether number of children affects wealth mobility at all among individuals in the NLSY.

While this subsection has shown that wealth inequality and racial-ethnic differences have been documented in detail, underlying processes behind the observed patterns have only recently been explored using high-quality longitudinal data. One exception is Painter's (2010) study of early labor force participation and later-life wealth accumulation suggests that adolescent employment helps individuals to develop certain kinds of human capital that are positively linked to educational attainment, which we might expect could in turn positively influence long-term wealth accumulation (this will be discussed in greater detail below). Specifically, by acquiring work experience and being exposed to various business environments, employed adolescents seem to develop practical life skills, knowledge, abilities, and resources that favorably shape educational attainment, career outcomes, adult financial decision making, and early social networks, all of which are processes that could correlate with later earning opportunities, investment decisions, and upward income and wealth mobility during adulthood.

## **Research Questions**

In an effort to modestly enhance what social scientists know about wealth dynamics among young baby boomers from the NLSY 1979 dataset, I begin by presenting descriptive data aspect of wealth in the U.S. that previously have received too little attention, namely wealth levels among persons in early mid-life:

- 1) What were the wealth levels of sampled individuals during their mid-20s and later during their mid-40s when they were entering their years of peak earnings capacity?
- 2) How much inequality in wealth—and what degree of racial inequality in wealth—existed at the start of and across the 20-year period when my sampled cohort of respondents aged from 24 to 44?

Then, the description turns to issues of intra-generational wealth dynamics:

- 3) What proportion of individuals experienced upward and downward wealth mobility during the 20-year period, and to what extent did wealth mobility vary by race-ethnicity?
- 4) How much volatility (upward or downward fluctuation) in people's wealth was there as they aged, and how did wealth volatility vary by race-ethnicity?

Finally, this study connects substantively to the results discussed previously in Section 3 by modeling wealth mobility in relation to timely college completion and upward wealth mobility in mid-adulthood, and then (following on other studies, e.g., Painter 2010) by analyzing potential variation in the influence of adolescent employment on upward wealth mobility:

- 5) People who complete college “on time” may have various traits that would make them more likely to accumulate wealth later on. So, does timely completion of four-year college predict upward wealth mobility in adulthood?
- 6) Employment is the most common transition into an “adult role” that teenagers undertake, and transitions into roles typically associated with adulthood may partially explain differences in wealth mobility. (Adolescent employment also is associated with behaviors and traits that portend greater chances of accumulating wealth during

adulthood.) So, to what extent does adolescent employment help to predict upward wealth mobility in adulthood?

## **Data and Methods**

To analyze within-person or “intra-generational” wealth mobility I selected respondents from the NLSY who were 24 years old in 1985 and who were interviewed in that year and also interviewed in 2004 when they were 44 years old. To be included in the analyses, these individuals also had to have been interviewed in at least two of the years between 1985 and 2004. I started at 1985 because that is the earliest year for which the best quality array of wealth variables was available (see Appendix D for discussion of this and for more detail about wealth in the NLSY). I defined the “later age” as ending at age 44 in part because I wanted to capture individuals’ pre-retirement lives (retirees tend to spend down their wealth or transfer it *inter vivos*, and they are more likely than pre-retirees to have inherited at least some wealth, either of which would distort the measurement of accumulated wealth). In measuring wealth toward the beginning of individuals’ working lives and then during their early peak working years, I capture the period when wealth tends to stabilize and peak (Conley 2007, 15). The measures used to capture intra-generational wealth mobility are described below.

## **Variables**

For descriptive analyses presented below, the basic outcomes of interest were amount of wealth and wealth quartile position. Initially I examined deciles but decided that it would be more parsimonious (and just as revealing) to use the quartile categorization. It may seem unusual or unwise to treat wealth as a relatively crude ordinal variable rather than a continuous one, but

there are several reasons for doing so. One is that, at certain points—say, when an individual moves from the ninth decile to the tenth (top) decile or from the fourth to the fifth decile—the movement is less meaningful than when an individual moves from the sixth decile into the eighth decile. Moving from ninth to the top merely indicates that a wealthy person got a bit wealthier, a shift that is unlikely to change a person’s life as meaningfully as if they had moved from the sixth to the eighth decile, a kind of movement that still gets captured using quartiles—moving from the sixth decile up to the eighth represents a similar thing as when one moves from the second quartile up to the first (top) quartile. Similarly, showing that a group of individuals moved from the second-lowest decile into the lowest decile would simply say they went from a very low-wealth position to a slightly lower-wealth position; arguably it would be as meaningful to capture whether individuals moved substantially (instead of only slightly) downward in the wealth distribution. A more general point is that at very low and very high levels of wealth, the actual amount of wealth people hold is not very relevant in mobility analyses; what is most of interest to researchers is people’s relative wealth positions compared to others. Another reason for using a “positional” measure of wealth is that wealth is very unstable compared to income. Indeed, one of the biggest problems wealth researchers have is that wealth can change dramatically year to year, and even weekly or monthly in ways that income does not. This may be due to stock market fluctuations, the selling/buying of a car or home, an expected or unexpected windfall, sudden emergencies that require a family to dip into its savings, and so forth. One other reason why researchers commonly use ordinal measures of wealth with this particular dataset is that although the actual measurement of people’s wealth in the NLSY is very detailed and includes many dimensions, it is also much less precise than it might first appear—when asked to estimate their net worth, a common response among NLSY respondents is “don’t

know,” in which cases the NLSY asked individuals to provide brackets indicating a range for the high/low estimate of their wealth during a given year. (This is very different from income—most people accurately know what they earn.)<sup>46</sup>

For these and other reasons, researchers (e.g., Conley 2008; Steckel and Krishnan 2011) have tended to use relatively coarse categorizations of wealth. In my analyses of wealth mobility presented below, I will focus on whether individuals had moved into or persisted within the top/bottom wealth quartiles as 44 year-olds given where they started as 24 year-olds. In the logit models for top-bottom wealth quartile mobility and persistence, the outcome variables were persistence within and movement out of or into the highest/lowest quartiles by age 44 given where in the distribution individuals started at age 24. First, and in order to remind the reader of what was at issue in the preceding subsections about wealth’s significance, the paragraphs below will briefly mention some of the scholarly literature in addition to the discussions of measurement.<sup>47</sup>

I included racial-ethnic status as an independent variable (black and Hispanic, with white as reference category), which simply will be referred to as “race-ethnicity” throughout.<sup>48</sup> Black and Hispanic Americans typically have lower wealth levels compared to whites, and part of my goal is to account for this disparity. For annual family income, I follow common practice by averaging family income across several early ages (24 to 30) and across later ages (38 to 44) in

---

<sup>46</sup> This point was clarified for me through a personal email correspondence with Dr. Jay Zagorski, the official handler of NLSY’s wealth variables at the Bureau of Labor Statistics.

<sup>47</sup> The yearly wealth variables summarize a great deal of information: how much people’s homes, cash savings, stock and bond portfolios, estates, businesses, and automobiles were worth, as well estimates of the value of people’s mortgage debt, property debt, and other debt. Creating the household wealth variables involved help from experts who manage the NLSY “wealth section” (which includes approximately 20 questions on assets). All wealth data were adjusted for inflation using constant 2000 dollars.

<sup>48</sup> As indicated in previous sections of this dissertation, the race-ethnicity measure for “black” is based on the NLSY interviewers’ judgments rather than respondents’ reported self-identifications. Also, as has been done in most other studies using the NLSY 1979 cohort, “Asian” and “American Indian” respondents were omitted due to substantial *n* limitations in the initial sample.

order to achieve a more stable measure of typical income (see, e.g., Haveman et al. 1991; Duncan et al. 1998; Gottschalk 1997; Torche 2011). Another explanatory variable is "family economic resources," by which I mean a measure of family poverty during respondents' mid- to late-adolescence (discussed below). Parental financial resources have long been known to affect children's academic performance and educational attainment (McLanahan and Sandefur 1994; Duncan 1994). Income in individuals' families-of-origin (during the time when there were living in their parents' household) can provide a reasonable measure of a household's economic status, and it also is more stable than earnings, so I followed the lead of other researchers (Mayer and Lopoo 2004) in assuming that family income when the respondents were aged 14 to 21 is strongly correlated with what their family-of-origin's income was when they were very young (birth to early teens). The experience of poverty, as one indicator of family economic resources, influences a family's ability to provide or opportunities, which could plausibly affect later educational attainment and economic well-being in adulthood. The experience of poverty in one's family-of-origin is commonly used as a relevant measure of family socioeconomic resources,<sup>49</sup> so I included a variable indicating whether respondents lived below the federal poverty line (adjusted to the consumer price index for each) for at least one year while they were living in their family-of-origin's household.

To further conceptualize the socioeconomic status of individuals' families of origin, I included a parental education variable for respondents' mothers' total years of formal schooling. Following other studies (e.g., Conley 2001), I included an indicator of the mother's years of

---

<sup>49</sup> As I mentioned previously in sections 2 and 3, this measure of family poverty may seem somewhat crude, but it differs from the "window variables" discussed in the dissertation's introduction (Section I) because families and individuals who had spent "at least one year below the poverty line" typically would also qualify as having been two or three years below poverty line, a meaningful correlation that hints at a kind of convergent validity among the latter and former variables. By contrast, problems of measurement validity are more pernicious with true window variables (e.g., the variable "lived with both biological parents at age 14" is not necessarily associated with the typical/modal type of parental presence during an individual's childhood years).

formal education. Based on conventional wisdom in mobility research, I expect parental educational attainment, especially that of mothers (they are the most consistently present parental figure across household types), to predict respondents' wealth positions and of their chances for wealth mobility in later life. Other research shows a correspondence between an individual's educational attainment and that of their eldest sibling's education, which has been theorized as a "role modeling effect" (Benin and Johnson 1984; Hauser and Wong 1989), so I have I included an indicator for eldest sibling's total years of formal schooling.

The presence of siblings represents an important dimension of an individual's family context. Parents' ability to share material and financial resources with a child will tend to diminish with more children in a household. Hence, is some evidence shows that number of siblings negatively influences children's educational attainment, in part because having more siblings can dilute family resources and cause relationship strain between parents and general stress within the household (Kuo and Hauser 1996). Given these reasons for expecting a negative effect of siblings on later wealth, in the logit models presented later I included a count of siblings indicated by the number of siblings a respondent ever had from birth to age 18.

Another explanatory variable, teenage childbearing, is defined as whether a respondent became a parent before turning 20 years old. There is a tremendous amount of scholarship on teenage parenthood, with researchers, policymakers, and the lay public decrying its apparent negative effect on various socioeconomic outcomes (Moore and Waite 1977; Furstenburg et al. 1989; Furstenburg 2003). While scholars have examined teen parenthood's effect on a person's socioeconomic mobility, but little work has examined that in the context of wealth .

The presence of children in the family during an individual's mid-life years may also affect wealth, although some uncertainly about its effect prevails among scholars as mentioned



previously. Having children is expensive pre- and post-natally, and their arrival may lead to financial strain. Alternatively, it is possible that the expectation or arrival of children might motivate parents to save money and acquire assets in anticipation of the expenses that accompany this life transition. Having children, and knowing what they will cost, might induce in parents a desire to begin accumulating wealth (e.g., through homeownership) earlier in life than they would have done otherwise (see Mulder 2006). Although there is still uncertainty as to the specific causal connections between number of children and wealth in later life (Painter and Shafer 2011), I expect my analyses will support the resource diminution hypothesis, that the presence of one or more children compels parents to spend down their wealth more than it induces parents to save. I have included as a predictor the number of children in individuals' households when respondents were in their 20s and 30s.

To account for another key aspect of “the transition to adulthood” discussed previously, I included a variable indicating whether individuals engaged in paid employment during any of their teenage years. Employment is by far the most common transition into an adult role that teenagers undertake, and transitions into roles typically associated with adulthood may partially explain differences in wealth mobility. Along with timely college completers, people who worked as teenagers may be more likely to have the kind of behaviors, human capital, and social connections that portend greater long-run chances of building wealth and holding onto it compared to those who did not work as teenagers.

### **Sample Statistics and Descriptive Analyses**

As a reminder, the initial descriptive analyses presented here are meant to enhance what social scientists know about wealth dynamics among young baby boomers in the NLSY 1979 dataset,

which is a relatively underexplored dataset when it comes to wealth. I begin by addressing the following basic questions:

- 1) What were the wealth levels of sampled individuals during their mid-20s and later during their mid-40s when they were entering their years of peak earnings capacity?
- 2) How much inequality in wealth—and what degree of racial inequality in wealth—existed at the start of and across the 20-year period?

So as to provide some basic background to the analyses that follow, it is worth noting the wealth levels that individuals in this young baby boomer cohort had in their mid-20s and later in their mid-40s, and it also is worth looking descriptively at the degree of wealth inequality at these ages. First I compare whites, blacks, and Hispanics on wealth levels and mobility as they aged.

Table 4.1 presents means and other descriptive statistics for individuals at age 24 in 1985 and when they were 44 in 2004. As might be expected, whites were much wealthier than blacks and Hispanics at each age, and over time they experienced a median wealth increase of slightly more than 700 percent, from \$6,050 to \$48,667. Blacks' and Hispanics' wealth also grew from the early to later age (by 600 and 564 percent, respectively), but the racial-ethnic wealth gaps widened. Blacks' and Hispanics' early median wealth was 50.4 and 82.6 percent of whites', respectively, and in their mid-40s it fell to 43.6 and 68.2 percent. This finding aligns with other analyses (e.g., Conley 2007) using different panel datasets over nearly identical ages—there were large percentage increases for all three racial-ethnic groups' net worth up to the mid-40s accompanied by a widening racial-ethnic wealth disparity.

**Table 4.1**  
**Sample Statistics by Race-Ethnicity When Individuals Were Aged 24 and 44**

Age 24:	Total Sample	White	Black	Hispanic
<b>Married (%)</b>	45.05	46.52	38.89	36.65
<b>Education (yrs)</b>	12.73	12.79	12.49	12.32
<b>Household Wealth</b>	11,280	12,240	6,962	9,084
<b>Median Household Wealth</b>	5,500	6,050	3,050	5,000
<b>Household Income</b>	16,108	16,536	14,188	14,542
<b>Family Poverty (%)</b>	10.70	9.83	14.62	19.79
<b>Mother's Education (median yrs)</b>	12	12	12	9
<b>Eldest Sib's Education (median yrs)</b>	12	12	12	12
<b># of Siblings During Childhood</b>	3.79	3.35	4.91	4.71
<hr/>				
Age 44:				
<b>Married (%)</b>	89.04	90.75	81.85	89.75
<b>Household Wealth</b>	67,921	79,007	35,865	49,348
<b>Median Household Wealth</b>	42,250	48,667	21,233	33,183
<b>Household Income</b>	69,511	73,862	56,251	62,930
<b>Family Poverty (%)</b>	5.34	4.64	6.70	7.22
<b>Education (median yrs)</b>	13.82	14	13	12
<b>Time to College (median yrs)</b>	23	22	24	24
<hr/>				
N =	2,093	1,437	336	320

Sample statistics are reported as means unless indicated otherwise. Income and wealth values are reported in constant 2000 dollars (adjusted using the CPI).

Descriptively it is revealing to examine issues of intra-generational wealth mobility, and here I do this by addressing two questions:

- 1) What proportion of individuals experienced upward and downward wealth mobility from age 24 to 44, and to what extent did wealth mobility vary by race-ethnicity?
- 2) How much volatility (upward or downward fluctuation) in people's wealth was there as they aged, and how did wealth volatility vary by race-ethnicity?

The persistence of low inter-generational mobility in relative wealth position for individuals, especially at the top and bottom ends of the U.S. wealth distribution, has become common knowledge among stratification researchers (e.g., Lusardi and Mitchell 2007). “Stickiness at the ends” is the conventional phrase connoting that it is much harder for people whose parents had very low wealth to move toward the top of the wealth ladder than it is for offspring of very wealthy parents to stay at the top. But since the NLSY 1979 data have only very limited information on respondents' parental wealth, I address several underexplored questions about intra-generational wealth mobility, or movement along the wealth distribution that individuals experienced during their lifetimes. First I interpret some descriptive results and address the question of intra-generational wealth mobility for the overall sample, and then turn to racial differences in wealth mobility. Later I will present results from mobility analyses with logit and ordered logit models, asking how it might be influenced by timely college completion and adolescent employment, among other factors.

#### *Correlations of Wealth: Emerging Adulthood to Mid-Adulthood*

Descriptively we can begin to understand mobility differences based on where in the wealth distribution one begins by looking at how early wealth correlates with later wealth. (From here

onward, the terms “bottom” and “top” shall refer to the lowest and highest quartiles containing the least wealthy and wealthiest individuals.) I begin with intra-generational correlations of wealth, that is, the degree to which wealth during early adulthood correlates with wealth in later adulthood. Table 4.2 indicates that, among everyone in the sample, wealth levels during early adulthood (when individuals were age 24) were correlated with wealth levels later on at only 0.24. Early wealth does not seem to account for much of where individuals ended up on the wealth ladder later on when they were 44 years old.

We know that wealth inequality persists along racial lines, but within a generation should we expect the gap between blacks and whites to narrow or widen over a two-decade period? Perhaps we should expect it to narrow considering the post-1960s increases in minorities’ access to employment, home loans, and public higher education. Or perhaps we should expect it to widen considering the covert effects of institutional discrimination in labor and housing markets and the lingering influence of a long history of inequality in access to wealth-building opportunities (chattel slavery, or discrimination in banking, or the practice of redlining, and so forth). The question of racial difference in intra-generational mobility is complicated and will be explored below, but for now it is worth examining racial differences in the correlation of wealth over time. Table 4.2 displays estimates for racial-ethnic categories, showing that the intra-generational correlation of wealth was more than fifty percent greater for minorities compared to whites (0.32 and 0.31 vs. 0.21); early wealth seems more closely linked with later wealth for blacks and Hispanics than for whites. Overall the correlations between early wealth and later wealth are still quite low across racial-ethnic groups, but (at least descriptively here) the lower correlation for whites indicates more room for upward or downward movement among whites than Hispanics or blacks.

As has been done in other studies of wealth mobility, and in the interest of discerning the “stability” of wealth over time—one basic measure of which is how much of individuals’ overall wealth trajectories can be captured just by looking at their wealth for any given age. To this end, the third row in Table 4.2 displays the intraclass correlation coefficient for wealth. One minus the intraclass correlation coefficient indicates how much of an individual’s 20-year wealth trajectory is not captured by just one particular age. For example, among whites roughly 90 percent of wealth status is not captured by one just one yearly measure of wealth. The table shows similarly low intraclass correlations for the entire sample and for each racial-ethnic group. Wealth is notoriously unstable across people’s lifetimes. This result, though by no means a definitive statement, implies that wealth is very volatile during individuals’ adulthoods.

**Table 4.2**  
**Correlations of Wealth and Income: Ages 24 to 44**

	Sample	Whites	Blacks	Hispanics
Intra-generational Wealth Correlation	0.24	0.21	0.32	0.31
Intra-generational Income Correlation	0.25	0.21	0.33	0.35
Intraclass Correlation (Wealth)	0.06	0.10	0.08	0.09
Intraclass Correlation (Income)	0.14	0.15	0.12	0.13

#### *Wealth Quartiles and Intra-generational Quartile Mobility*

In descriptive analyses not shown, individuals in this sample experienced substantial wealth gains and losses from earlier to later ages, but such changes are entirely distinct from changes of individuals’ positions in the wealth distribution over time, a topic to which I now turn. Table 4.3 presents sample statistics for wealth quartiles, broken down by respondent's age. From the early

to later ages, median wealth for all quartiles increased with some minor differences in the percentage increase across quartiles (e.g., about a fourfold increase for the bottom quartile compared to a threefold increase for the third quartile; and just under a fourfold increase for the second quartile compared to more than fivefold increase for the top quartile). Overall the picture is what one would expect theoretically and in the context of prior research on wealth, namely that the members of each quartile experienced a several-fold percentage increase in their mean and median wealth as they aged.

**Table 4.3**  
**Sample Statistics for Wealth Quartiles, Ages 24 and 44 (constant 2000 dollars)**

	Mean	Median	Minimum	Maximum
<b>Quartiles: 24-yr olds</b>				
Bottom	87.20	990	-69,998	2,170
Third	4,086	4,000	2,200	6,200
Second	9,129	8,900	6,300	13,150
Top	29,877	21,500	13,200	178,500
<b>Quartiles: 44-yr olds</b>				
Bottom	3,261	4,000	230,000	8,000
Third	12,937	12,000	8,005	19,000
Second	30,798	29,000	19,200	50,000
Top	174,755	108,000	50,200	2,409,714

Table 4.4 is a descriptive mobility table showing movement of persons moving across wealth quartiles positions from when they were age 24 to age 44. First, over a quarter (28 percent) of individuals who started in the top wealth quartile remained in the top two decades later. Extreme upward and downward mobility were unlikely; less than a tenth (8.8 percent) of those who started in the bottom wealth quartile as 24 year-olds were in the top quartile later on.

More than a quarter (28 percent) of those who began at the top stayed there, only 17.4 percent of them fell into the bottom quartile, while nearly half (44.7 percent) of individuals who began in the bottom quartile remained there. These findings are in line prior research using other panel datasets—they indicate “stickiness at the ends,” or relative immobility among people who began in the highest and lowest quartiles of the wealth distribution. At least according to this descriptive picture, as people age it seems very difficult for initially low-wealth individuals to move into the top quartile—a vast majority of these them never made it to the top and a substantial plurality stayed at the bottom. Especially striking is how this contrasts with high-wealth individuals, who seemed better able to hold on to their wealth; a majority of these individuals moved downward in the distribution, but more than one quarter (28 percent) were at the top as 44 year-olds.

**Table 4.4**  
**Description of Intra-generational Quartile Mobility in Wealth (percentages)**

Quartile Position at Age 24	Quartile Position at Age 44			
	Bottom	Third	Second	Top
<u>Overall Sample</u>				
Bottom	44.7	28.2	18.2	8.8
Third	29.4	29.3	28.1	13.1
Second	21.2	29.5	29.4	19.9
Top	17.4	28.9	25.7	28.0
Total	29.1	28.9	25.1	16.8
<u>Whites</u>				
Bottom	41.5	26.4	20.8	11.3
Third	26.3	25.2	33.4	15.5
Second	20.9	25.7	30.5	22.9
Top	18.8	28.7	23.8	28.7
Total	27.0	26.5	27.0	19.5
<u>Blacks</u>				
Bottom	53.9	33.3	10.3	2.6



Third	33.3	37.5	16.7	12.5
Second	22.7	45.5	22.7	9.1
Top	12.5	18.8	31.3	37.5
Total	35.6	34.7	17.8	11.9

#### Hispanics

Bottom	44.0	28.0	20.0	8.0
Third	36.7	36.6	20.1	6.7
Second	21.1	31.6	31.7	15.8
Top	13.3	40.0	33.3	13.4
Total	31.5	33.7	24.7	10.1

---

As displayed in the upper panel of Table 4.4, there was relatively low top-bottom and bottom-top mobility in the sample overall. However, looking at mobility among the middle quartiles we find that wealth *mobility* seems to characterize this sample (this was hinted at earlier by the relatively low intra-generational and intraclass correlations for wealth shown in Table 4.2). There is substantial upward/downward mobility from the bottom and top quartiles into the third and second quartiles; although 44.7 percent of individuals who started in the bottom stayed there, nearly half (46.4 percent) of them moved upward into the third or second quartile. And among those who started in the third quartile (second-to-bottom), 41.2 percent of them moved upward into the second or top quartiles, and 58.7 percent remained in the bottom half (third or bottom quartile). Compared to those who began at the bottom, there was greater upward mobility for people who started out in this third quartile, as 13.1 percent of them ended up in the top bracket.

Among those who started in the second quartile (second-from-top), a fifth of them (19.9 percent) were able to move up to the top bracket. People in this second-highest quartile also experienced substantial *downward* mobility—as 44 year-olds, half (50.7 percent) of them were either in the third or bottom quartile. Finally, the general picture of wealth mobility also pertains to those who started at the top, as a much higher percentage of those in this top wealth quartile

moved downward than stayed at the top (54.6 percent moved downward to the second and third quartiles, 28.0 percent remained at the top).

### *Racial Differences in Intra-generational Wealth Mobility*

The three bottom panels in Table 4.4 display wealth quartile movement for whites, blacks, and Hispanics. A substantial proportion of whites and Hispanics (41.5 and 44 percent, respectively) who were in the bottom quartile as 24 year-olds remained there later on. By contrast, the corresponding figure for blacks was 53.9 percent. As well, among blacks and Hispanics who experienced any upward wealth mobility, only about a tenth of them (11.9 and 10.1 percent, respectively) made it to the top quartile. Among blacks, only 12.9 percent moved from the bottom quartile to the top half of the distribution, whereas 28 percent of Hispanics did so. By contrast, nearly a third of initially low-wealth whites (32.1 percent) moved into the top half of the distribution. The story here is that low-wealth whites experienced greater upward mobility and low-wealth blacks experienced the least amount of upward mobility.

Racial differences in downward mobility also appear in these tables.. For whites who were at the top of the distribution during their mid-20s, 28.7 percent remained there during their mid-40s, but only 13.4 percent of Hispanics did so. Somewhat unexpectedly, more than a third (37.5 percent) of top-quartile young blacks were able to remain there as 44 year-olds. Compared to blacks, Hispanics and whites who were at the top in early adulthood were less likely to retain their position over time. A higher percentage of whites and Hispanics fell from the top to second and third quartiles compared to blacks.

## **Logit Regression Models of Wealth Persistence and Mobility**

The previous discussion described wealth position at given ages, volatility in wealth over time, the share of individuals who experienced mobility (upward and downward), and how these were patterned along racial-ethnic lines. Here I use statistical models to present a clearer picture of how various factors may influence wealth mobility. Previously I discussed two other research questions motivating this work, and will briefly restate them here. First, people who complete college “on time” may have various personal traits or social characteristics that would make them more likely to accumulate wealth later on. They may be more predisposed to a long-run future orientation, or they may be favored by employers in hiring and promotion decisions, and so forth. So I wondered whether timely completion of four-year college predicts upward or downward wealth mobility in adulthood, controlling for where in the distribution individuals started out when they were 24. Secondly, employment is the most common transition into an “adult role” that teenagers undertake, and as I theorize below, transitions into roles typically associated with adulthood may help to account for differences in wealth mobility. (As with timely college completion, adolescent employment also is associated with behaviors and traits that portend greater chances of accumulating wealth during adulthood.) So, to what extent does adolescent employment help predict upward or downward wealth mobility in adulthood? In short, this subsection represents my attempt to model wealth mobility, specifically the association between timely college completion and wealth mobility in mid-adulthood, and the potential influence of adolescent employment on wealth mobility.

### *Adolescent Employment, Timely College Completion, and Wealth Mobility*

Teenage parenthood is often (though certainly not always or in all U.S. cultural communities) understood as a kind of precocity that should be avoided and admonished. As we have seen here and previously in section 3, teen parenthood is associated with dampened life chances, at least in terms of timely educational completion and later wealth position. But it is worth asking how other types of adolescent behaviors that adults tend to *encourage* but that are also widely seen as adult roles, like employment, might contribute to our understanding where people end up in the wealth distribution. Much research has documented the connection between adolescent employment and adult labor market outcomes such as income and wages, but despite extensive scholarship in that area (e.g., Bacolod and Holtz 2006; Hotz et al. 2002), little research has examined adolescent employment's effect on later-life wealth. Painter's (2010) recent work provides some important direction in this regard, and has suggested that adolescent employment does seem to improve people's net worth and overall financial well-being later in their lives.

Given this, and for theoretical reasons outline previously, I expect at least some causal connection between adolescent employment and wealth mobility in later life. Again, this line of reasoning is that employment, often associated with 'positive' consequences (rewards such as income, an implicit reassurance of upward occupational mobility, long-run financial stability, etc.) and financially beneficial personality traits, might indicate individuals' predilection to delay other kinds of short-term gains for eventual longer-term rewards. And indeed, recent research has upended the old perceptions that adolescent employment causes poor academic outcomes, delinquency, substance use, and other "negative" behaviors (e.g. Apel et al. 2006; Paternoster et al. 2003; Warren et al. 2000; Schoenhals et al. 1998). Rauscher (2011) found that adolescent employment increases people's likelihood of experiencing early transitions to fertility and

marriage, but not specifically to teenage parenthood. With this in mind, it is reasonable to posit early employment experience as an important indicator of the human capital resources that can help people get ahead socioeconomically throughout their lives.

There are no comprehensive theories of wealth mobility to guide a statistical model, so I follow others (e.g., Steckel and Krishnan 2006) and chose regressors based on recent theorizing and research about wealth accumulation instead of mobility. To recapitulate with respect to a couple of the key explanatory variables, I suspected teenage parenthood would be unfavorably associated with individuals' chances of moving upward into the top wealth quartile and that it would be associated with an increase in the chances that low-wealth 24 year-olds would remain in the lowest wealth position as 44 year-olds. As well, I expected that adolescent labor force participation would improve chances for upward wealth mobility among those who started out low-wealth as 24 year-olds. Conversely, for people who started out in the highest-wealth position, I expected a less important role for adolescent employment in terms of people's chances of holding onto wealth (presumably, factors other than having a job during adolescence are more powerfully influential for high-wealth individuals' capacity to hold onto their wealth over time). Other expectations about the data, discussed previously, will be mentioned below.

I followed other scholars (e.g. Steckel and Krishnan 2006) by focusing on mobility near the extremes of the wealth distribution , and specifically analyzed as outcomes people's movement into or persistence within the top and bottom wealth quartiles. To do this, I estimated four logit models, separated by initial position and later destination. I coded the dependent variable for staying in the top quartile as 1 if the individual was in the top quartile at both ages and 0 if the individual was in the top quartile at age 24 but not at age 44. Similarly, the dependent variable for movement into the top quartile = 1 if the individual was in the top quartile

at age 44 but not age 24, and = 0 if the individual was not in the top quartile at either age. The dependent variable for staying in bottom quartile = 1 if the individual was in the bottom quartile at both ages and = 0 if the individual was in the bottom quartile at age 24 but not at age 44. The dependent variable for moving into bottom quartile = 1 if the individual was in the bottom quartile at age 44 but not age 24, and = 0 if individual was not in the bottom quartile at either age.

Table 4.5 displays several notable results about persistence in and movement to the top/bottom quartiles of the wealth distribution. The first I will mention is that individuals who were parents as teenagers had much lower chances of moving into the top wealth quartile (0.295) relative to those who were not teenage parents, and they were slightly more likely ( $OR = 1.185$ ) fall into the bottom quartile. This was unsurprising; as we saw in this study's descriptive results and earlier in Section 3, teen parenthood may dampen life chances, at least in terms of timely educational completion and mid-life wealth position.

**Table 4.5**  
**Logit Models of Wealth Persistence in and Movement into the Top/Bottom Quartiles**

Variable	Staying in Top			Moving into Top		
	OR	Std. Err.	95% C.I.	OR	Std. Err.	95% C.I.
Timely College	1.832**	0.629	0.677-1.954	2.843**	1.611	1.524-3.294
Adolescent Employ.	0.931*	0.517	0.137-0.988	1.849*	0.626	0.486-1.973
Black	0.676*	0.377	0.529-0.981	0.579*	0.301	0.209-0.961
Hispanic	0.772*	0.655	0.146-0.792	0.955	0.499	0.342-1.664
Family Poverty	1.079	0.919	0.637-1.987	0.734*	0.347	0.491-1.185
Teen Parent	1.492*	0.721	0.319-1.916	0.295*	0.309	0.137-0.437
Mother's Edu.	1.200*	0.120	0.986-1.459	1.065	0.657	0.944-1.202
Eldest Sibling's Edu.	1.048	0.384	0.974-1.126	1.078*	0.767	0.937-1.239
Number of Siblings	0.810	0.114	0.615-1.067	0.905*	0.084	0.754-1.086
Lived w/ Both Parents	2.081**	0.802	1.671-2.459	0.501*	0.169	0.258-0.974
Number of children	0.551*	0.124	0.353-0.857	0.899	0.119	0.693-1.167
Sample Size	335			672		

	Staying at Bottom			Moving into Bottom		
	OR	Std. Err.	95% C.I.	OR	Std. Err.	95% C.I.
Timely College	0.188**	0.107	0.062-0.572	0.545**	0.153	0.313-0.947
Adolescent Employ.	0.248**	0.141	0.081-0.756	0.674*	0.197	0.381-1.194
Black	1.229*	0.385	0.407-1.920	1.249*	0.307	0.771-1.943
Hispanic	0.889	0.409	0.277-1.180	0.932	0.182	0.359-1.415
Family Poverty	2.085	0.879	0.911-3.767	1.302	0.295	0.834-2.030
Teen Parent	0.929*	0.406	0.245-2.170	1.185*	0.348	0.665-1.919
Mother's Edu.	1.079	0.756	0.941-1.238	0.994	0.034	0.870-1.004
Eldest Sibling's Edu.	0.981	0.084	0.844-1.139	0.951	0.037	0.841-0.986
Number of Siblings	1.085*	0.084	0.931-1.263	1.092*	0.046	1.004-1.185
Lived w/ Both Parents	1.078	0.657	0.864-1.675	1.017	0.228	0.739-1.661
Number of children	1.112*	0.326	0.626-1.976	1.148*	0.152	0.755-1.361
Sample Size	628			458		

\*p<0.05, \*\*p<0.01

Employment during adolescence, which in the stratification literature has been linked to various “positive” consequences (the reward of income, a psychological reassurance of upward occupational mobility, the prospect of long-run economic stability, etc.) and financially beneficial personality traits, may also indicate individuals’ predilection to delay other kinds of short-term gains for eventual longer-term rewards. On this basis I characterized early employment experience as an important indicator of the human capital resources that aid people in getting ahead socioeconomically, and I also posited that it would be favorably associated with

quartile mobility at the top and bottom of the wealth distribution. The results shown in Table 4.5 seem to support this expectation, as adolescent employment was positively associated with people's odds of upward mobility into the top quartile ( $OR=1.849$ ), and it was associated with only slightly reduced odds of persistence in the top ( $OR=0.931$ ). Controlling for other factors, persons who had worked when they were teenagers were much less likely to persist in the lowest quartile or to move into it ( $OR=0.248$  and  $0.674$ , respectively), suggesting a substantial beneficial effect of adolescent employment on wealth.

Previously I suggested that people who completed college in a “timely” manner (prior to age 24) could have other social characteristics or personality traits that would make them more likely to accumulate wealth later on; they may be more predisposed to a long run future orientation, more likely to delay short-term rewards for longer-term stability, or they may be favored more by employers in hiring and promotion decisions during their very early working lives compared to workers who had not completed by age 24. (Considering how commonplace educational endogamy is in the United States, it is also possible that timely completion works indirectly through the capacity of college-educated married couples to pool their wealth.) With this in mind, perhaps the most striking result from Table 4.5 is that those who were “timely” college completers—those who completed prior to age 24—had vastly better odds of staying in and of moving into the top quartile ( $OR = 1.832$  and  $2.843$ , respectively). And on the other end of the distribution, completing college prior to age 24 also seemed to have a negative (i.e., favorable) effect on people's chances of staying in or moving downward to the bottom quartile ( $OR = 0.188$  and  $0.545$ ).



## Concluding Discussion and Future Work

As a basic matter, researchers have not paid as much attention to the NLSY's detailed wealth data as one would expect, especially given this dataset's widespread use among social scientists who study socioeconomic inequalities (by far, NLSY primarily has been used in studies focusing on labor force participation and income inequality). Thus one contribution of my work here is that it is one of the few existing examinations of basic wealth dynamics among members of the NLSY 1979. Theoretically, a key idea that has been threaded throughout this dissertation has been that the *timing* of life events matters in various ways beyond their simple occurrence, especially when we consider events that evoke the "biographical" or individual notion of time as discussed much earlier in the introduction (Section 1). This assumption has long been standard in the life course tradition (Elder et al. 2004), but it has become elaborated theoretically and empirically as researchers uncover ever more complex phenomena related to when events happen in people's lives.

Indeed, much recent research has studied the sequencing of life events, particularly those occurring along the so-called "transition to adulthood," and the consensus is that life course trajectories among young people born in the United States are growing ever more variable in their occurrence and significance (O'Rand 2000; Roksa and Velez 2012). A recent estimate suggests that only about a quarter of young adults today follow the traditional chronological sequence of schooling, full-time job, getting married, and having children (Mouw 2005). One simple explanation here may be class-related fertility differentials, which reflects a change in the background circumstances of the population "at risk" of non-standard trajectories. As well, the expansion and differentiation of higher education over the last century both seem to have altered the meaning and timing of school-to-work transitions (Kerckhoff 2002), which could also have

implications for how researchers study the connection between “timely” college completion, adolescent or early-adulthood employment, and wealth mobility. If life course transitions among post-Baby Boom generations are becoming “demographically dense” (Rindfuss 1991), with multiple transitions occurring in a short amount of time, and if many young adults entering higher education cannot focus exclusively on their studies but instead combine their schooling with transitions into other social roles (Roksa and Velez 2012), then we would expect different results among these younger cohorts when studying the seeming positive influence of adolescent employment on adult wealth. (Of course, examining this issue as it pertains to members of very recent generations would require using a dataset other than the NLSY 1979.) As well, in the current era of substantially increased college prices, employment during college has understandably become more standard in the population.

One potentially important factor left out of consideration here is the role of academic achievement in wealth mobility, and more particularly academic achievement’s relation to racial disparities in wealth. Scholars have long observed a gap in academic achievement (e.g., scores on math and reading tests) between Black and white students, which apparently begins prior to kindergarten and persists into adulthood (Phillips et al. 1998; Campbell et al. 1999; Jencks and Phillips 1998; Yeung and Conley 2008). This early ability gap may intervene in the apparent connection between race-ethnicity and people’s chances of building wealth in adulthood. Since I did not account for differences in cognitive ability or specific academic skills, there is no way to tell the extent to which these might play a role in predicting where people end up on the wealth ladder as adults. It makes sense, then, to suggest (for my future work and perhaps that of other researchers) looking into individual-level abilities that develop early in people’s lives. However,

a different dataset would be needed for this since the NLSY does not contain retrospective data on what people's cognitive skills or academic abilities were in their very early lives.

The reader will note that I have not included gender in my analyses of wealth. In the NLSY, separating wealth by gender is conceptually difficult, since married individuals will include assets held by a spouse when asked about their wealth. In statistics not shown here, within the 24-44 age range women consistently had slightly higher average wealth than men, with women's mean wealth across these ages ranging from 5 to 7 percent higher than men's. However, medians tell a different story than means; women's median wealth was consistently *lower* (though only slightly lower) than men's median wealth during these years. Married women reported higher amounts of median wealth than married men, and unmarried men reported higher median wealth than unmarried women. There is a plausible explanation for this; as has been shown in other analyses of NLSY wealth data, these median calculations most likely are picking up the large number of unmarried, relatively poor women, which pull the median downward, while the mean calculations are picking up the much smaller number of relatively rich married women, which pushes the mean upward).<sup>50</sup>

Another admittedly daunting issue that I plan to address in the near future is how to sensibly account for individuals' relationship histories. Zagorsky (2005) has shown that, over time, single respondents increase their net worth very slowly, and by contrast the wealth of married individuals tends to increase, on average, by about 16 percent for each year of marriage. As well, Zagorsky showed that people will experience an average wealth drop of 77 percent after they divorce. While in percentage terms divorce hurts women more than men, the absolute difference is relatively small in the United States. Marital histories and cohabitation undoubtedly

---

<sup>50</sup> This phenomenon was first revealed in Zagorsky (1999).

will affect people's reported net worth, so this seems important thing to include in future work on wealth in the NLSY 1979. Transitions into and out of marriage (and when these events occur) surely will influence people's wealth trajectories, so it would be worthwhile to track people's relationship histories. It might also be useful to get details on the "terms" of people's divorces and separations. With this to mind, future researchers might consider using the detailed marital and relationship histories available in NLSY to try and detect whether fluctuations in reported wealth among men and women are accounted for by changes in marital or relationship status. This may provide a better picture of the way wealth is unequal by gender and thus would help researchers who pose questions about wealth accumulation or mobility among men and women over time.

Another possible direction for future work is indicated by the connections among a couple variables already examined herein. Specifically in addition to the reasoning about adolescent employment previously discussed, there is theoretical justification to suspect that early life behaviors such as adolescent employment, which is seen (and, arguably, felt and experienced) as an "adult" role, could be causally related to or correlated with other ones such as adolescent fertility. As I mentioned previously, recent research challenges the long-held perceptions that adolescent employment causes poor academic outcomes, delinquency, and substance use. However, adolescent employment may still put youth at higher risks for problematic behaviors (Bachman & Schulenberg, 1993; Brame, Bushway, Paternoster, & Apel, 2004; Ploeger, 1997), in part because it places them in social environments (workplaces) where they are likely to be in regular contact with adults, adult behaviors, and adult norms regarding what one should seek in life. An avenue for future research is to model the causal connections among kinds of teenage employment, perhaps even examining whether *early* teenage

employment (employment at ages 15 or 16 compared to ages 18 or 19) more negatively or positively influences wealth mobility later in life. The timing of adolescent employment may matter as much or more than the simple fact of it, and testing for these potential effects would be an interesting challenge.

It also would be intriguing to pose a more refined hypothesis about adolescent employment, especially in terms of its intensity and duration, and how it affects wealth at mid-adulthood. Surely just “having a job” cannot in itself ensure that employed adolescents will acquire the human capital positively associated with their adult wealth position. There may be variability in *kinds* or *durations* of employment, with some kinds or durations mattering more than others in their effects on later wealth—it is plausible to suspect that young people with limited or sporadic employment may receive some human capital benefits, but those with more intense work experiences or generally greater work investments may have better opportunities to build their human capital (or they may be more likely to work during their college years, which could portend greater wealth mobility by mid-adulthood). As a recent study by Painter (2010) showed, the intensity and duration of employment during adolescence seems to have effects on people’s later wealth positions, with people whose adolescent work experiences were of longer duration experiencing greater wealth accumulation in adulthood. Thus, examining people’s adolescent employment experiences by using a typology of intensity and duration (high-intensity/duration, low-intensity/duration, etc.) would seem a logical next step in clarifying what appears to be a positive effect of early employment on people’s wealth mobility into mid-adulthood.

## REFERENCES

- Aronson, M. and Michael Lacy. 2014. "Propensity Scoring and Other Methods for Decomposing Direct and Indirect Effects in Binary Response Models." Presented at the Midwest Sociological Society Annual Meetings, Omaha, NE. April.
- Bahr, P. 2009. "Educational Attainment as Process: Using Hierarchical Discrete-Time Event History Analysis to Model Rate of Progress." *Research in Higher Education* 50: 691-714.
- Barro, S., and A. Kolstad. 1987. *Who Drops Out of High School? Findings from High School and Beyond*. Washington, D.C.: National Center for Education Statistics (CS 87-397c).
- Becker, G. 1975. *Human Capital*. New York: Columbia Press.
- Behnke, A., K. Piercy, and Marcelo Diversi. 2004. "Educational and Occupational Aspirations of Latino Youth and Their Parents." *Hispanic Journal of Behavioral Sciences* 26(1): 16-35.
- Bernhardt, E. 1993. "Fertility and Employment." *European Sociological Review* 9:25-42.
- Blau, P. and O. Duncan. 1967. *The American Occupational Structure*. New York: Wiley.
- Blau, D. 1999. "The Effect of Income on Child Development." *The Review of Economics and Statistics* 81(2): 261-276.
- Blake, J. 1981. "Family Size and the Quality of Children." *Demography* 18: 421-442.
- Bowles, S. and H. Gintis. 2001. "The Inheritance of Economic Status: Education, Class and Genetics." Working Paper, *The Santa Fe Institute*.
- Bowles, S., H. Gintis, and Melissa Groves (eds.). 2005. *Unequal Chances: Family Background and Economic Success*. Princeton: Princeton University Press.
- Box-Steffensmeier, J. and B. Jones. 2004. *Event History Modeling: A Guide For Social Scientists*. New York: Cambridge University Press.
- Breen, R., K. Karlson, and Anders Holm. 2013. "Total, Direct, and Indirect Effects in Logit and Probit Models." *Sociological Methods & Research* 42(2):164-191.
- Brewster, K. and R. Rindfuss. 2000. "Fertility and Women's Employment in Industrialized Nations." *Annual Review of Sociology* 26: 271-296.
- Bumpass, L. and S. McLanahan. 1989. "Unmarried Motherhood: Recent Trends, Composition, and Black-White Differences." *Demography* 26(2): 279-286.

- Cameron, S. and James Heckman. 1991. "The Dynamics of Educational Attainment for Blacks, Whites and Hispanics." Working Paper. Chicago IL: University of Chicago, September 1991
- Chapman, C., J. Laird, and Angelina KewalRamani. 2011. "Trends in High School Dropout and Completion Rates in the U.S., 1972-2008." U.S. Department of Education: *National Center for Education Statistics*.
- Chen, Z. and H. Kaplan. 1999. "Explaining the Impact of Family Structure during Adolescence on Adult Educational Attainment: A Longitudinal Study." *Applied Behavioral Science Review* 7(1): 23-40.
- Cleves, M., W. Gould, R. Gutierrez, and Yulia Marchenko. 2008. *An Introduction to Survival Analysis Using Stata, 2<sup>nd</sup> Edition*. Stata Press.
- Conger, R., X. Ge, G. Elder, F. Lorenz, and R. Simons. 1994. "Economic Stress, Coercive Family Process, and Developmental Problems of Adolescents." *Child Development* 65: 541-561.
- Collins, Randall. 1990. *Weberian Sociological Theory*. Cambridge: Cambridge University Press.
- Conley, D. 2001. "Capital for College: Parental Assets and Postsecondary Schooling." *Sociology of Education* 74(1) 59-72
- Conley, D. 2000. "Sibship Sex Composition: Effects on Educational Attainment." *Social Science Research* 29: 441-457.
- Conley, D., K. Strully, and Neil Bennett. 2003. *The Starting Gate: Birth Weight and Life Chances*. Berkeley, CA: UC Press.
- Cooksey, E. 1997. "Consequences of Young Mothers' Marital Histories for Children's Cognitive Development." *Journal of Marriage and the Family* 59,2: 245-261.
- Dalaker, J. 2001. *Poverty in the United States: 2000*. US Census Bureau, Current Population Reports, Series P60-214. Washington, D.C.: US Government Printing Office.
- Dourleijn, E., A. Liefbroer, and Gijs Beets. 2000. "The Measurement of Educational Attainment in the FFS: Comparing the ISCED-classification with Information from Educational Histories in 17 European Countries." Paper prepared for the FFS Flagship Conference, Brussels, May 29.
- Duncan, G.J., and J. Brooks-Gunn. (eds.) 1997. *The Consequences of Growing up Poor*. New York: Russell Sage Foundation.

- Duncan, G., W.J. Yeung, J. Brooks-Gunn, and Judith Smith. 1998. "How Much Does Childhood Poverty Affect the Life Chances of Children?" *American Sociological Review* 63:406-423.
- Duncan, G., Kalil, A., Mayer, S., Tepper, R., and M. Payne. 2002. "The Apple Does Not Fall Far From the Tree." Institute for Policy Research: Working Paper WP-02-17. Institute for Policy Research, Northwestern University.
- Elder, G., M. Johnson and Robert Crosnoe. 2003. "The Emergence and Development of Life Course Theory" in Mortimer, J. and Michael Shanahan (eds.) *Handbook of the Life Course*. Springer Press.
- Elman, C., and A. O'Rand. 2004. "The Race is to the Swift: Socioeconomic Origins, Adult Education, and Wage Attainment." *American Journal of Sociology* 110(1): 123-160.
- Escamilla, S. and J. Santhiveeran. 2005. "Power of Teen Motherhood in Predicting Later Educational Attainment." *The Social Policy Journal* 4,1: 65-79
- Esping-Andersen, Gosta. "Untying the Gordian Knot of Social Inheritance." *Research in Social Stratification and Mobility* 21:115-138.
- Faherty, V. 2006. "Social Welfare Before the Elizabethan Poor Laws: The Early Christian Tradition, AD 33 to 313." *Journal of Sociology & Social Welfare* 33(2):107-122.
- Featherman, D., and Robert Hauser. 1978. *Opportunity and Change*. New York: Academic Press.
- Fomby, P. and A. Cherlin. 2007. "Family Instability and Child Well-Being." *American Sociological Review* 72 (April: 181-204).
- Francesconi, M. 2007. "Adult Outcomes for Children of Teenage Mothers." Institute for the Study of Labor (IZA), Discussion Paper No. 2778.
- Frase, M.J. 1989. "Dropout Rates in the United States: 1988." Washington, DC: National Center for Educational Statistics.
- Frisco, M. 2008. "Adolescents' Sexual Behavior and Academic Attainment." *Sociology of Education* 81: 284-311.



- Furstenberg, F. 1987. "Race Differences in Teenage Sexuality, Pregnancy, and Adolescent Childbearing." *The Milbank Quarterly* 65(2): 381-403.
- Furstenberg, F., J. Brooks-Gunn, and Lindsay Chase-Lansdale. 1989. "Teenaged pregnancy and childbearing." *American Psychologist* 44(2): 313-320
- Furstenberg, F. 2003. "Reflections on the Future of the Life Course." *Handbook of the Life Course*, Jeylan Mortimer and Michael Shanahan, eds. Springer.
- Gamoran, A. 1992. "The Variable Effects of High School Tracking." *American Sociological Review* 57: 812-828.
- Garasky, S. 1995. "Effects of Family Structure on Educational Attainment: Do the Effects Vary by Age of the Child?" *American Journal of Economics and Sociology* 54,1: 89-104.
- Geronimus, A.T. 2003. "Damned if You Do: Culture, Identity, Privilege, and Teenage Childbearing in the United States." *Social Science & Medicine* 57(5): 881-893.
- Geronimus, A.T., J. Bound, and T.A. Waidmann. 1999. "Health inequality and population variation in fertility-timing." *Social Science & Medicine* 49(12): 1623-1636
- Goldenberg, C., R. Gallimore, L. Reese, and Helen Garnier. 2001. "Cause or Effect? A Longitudinal Study of Immigrant Latino Parents' Aspirations and Expectations, and their Children's School Performance." *American Educational Research Journal* 38(3): 547-582.
- Goldrick-Rab, S. 2006. "Following Their Every Move: An Investigation of Social Class Differences in College Pathways." *Sociology of Education* 79(1): 61-79.
- Goldschmidt, P. and Jia Wan. 1999. "When Can Schools Affect Dropout Behavior? A Longitudinal Multilevel Analysis." *American Educational Research Journal* 36:715-738.
- Gottfried, A. 1984. *Home Environment and Early Cognitive Development: Longitudinal Research*. Orlando, FL: Academic Press.
- Gottschalk, P. "Inequality, Income Growth, and Mobility: The Basic Facts." *Journal of Economic Perspectives* 11(2):21-40.
- Grubb, W. 1993. "The Varied Economic Returns to Post-Secondary Education: New Evidence from the Class of 1972." *The Journal of Human Resources* 28: 365-382.
- Grusky, D. and T. DiPrete. 1990. "Recent Trends in the Process of Stratification." *Demography* 27:617-637.

- Hansen, T., S. McLanahan, and E. Thomson. 1997. "Economic Resources, Parental Practices, and Children's Well-Being." pp. 190-238 in *Consequences of Growing Up Poor*, Duncan, G. and J. Brooks-Gunn (eds.) New York: Russell Sage Foundation.
- Harding, D., C. Jencks, L. Lopoo, and Susan Mayer. 2005. "The Changing Effect of Family Background on the Incomes of American Adults." In Bowles, Gintis, and Groves (eds.). *Unequal Chances: Family Background and Economic Success*.
- Haas, S. and N. Fosse. 2008. "Health and the Educational Attainment of Adolescents: Evidence from the NLSY97." *Journal of Health and Social Behavior* 49(2):178-192.
- Haveman, R., B. Wolfe, and K. Wilson. 1997. "Childhood Poverty and Adolescent Schooling and Fertility Outcomes: Reduced-Form and Structural Estimates." pp. 419-460 in *Consequences of Growing Up Poor*, Duncan, G. and J. Brooks-Gunn (eds.) New York: Russell Sage Foundation.
- Hauser, R. 1999. *Should We End Social Promotion: Truth and Consequences*. CDE Working Paper No. 99-06. Madison, WI: Center for Demography and Ecology at UW-Madison.
- Heckman, J. and Yona Rubinstein. 2001. "The Importance of Non-Cognitive Skills: Lessons from the GED Testing Program." *The American Economic Review* 91(2) 145-149.
- Hetherington, E. M., Bridges, M., & Insabella, G. 1998. "What matters? What does not? Five perspectives on the association between marital transitions and children's adjustment." *American Psychologist* 53(2): 167-184.
- Hill, M., and Greg Duncan. 1987. "Parental Family Income and the Socioeconomic Attainment of Children." *Social Science Research* 16:39-73.
- Hoffman, S. 1998. "Teenage Childbearing is Not So Bad After All...Or is it? A Review of the New Literature." *Family Planning Perspectives* 30: 236-239.
- Hoffman, S., E.M. Foster, and F. Furstenberg Jr. 1993. "Reevaluating the Costs of Teenage Childbearing." *Demography* 30: 1-13.
- House, J. S., Lantz, P. M., & Herd, P. 2005. "Continuity and change in the social stratification of aging and health over the life course: Evidence from a nationally representative longitudinal study." Americans' Changing Lives Study. *Journal of Gerontology-SS*, 60B (Special Issue II), 15-26.

- Hout, M. and D. Dohan. 1996. "Two Paths to Educational Opportunity: Class and Educational Selection in Sweden and the United States." Pp. 207-231 in *Can Education Be Equalized? The Swedish Case in Comparative Perspective*. Erikson, R. and J. Jonsson (eds.). Boulder: Westview Press.
- Heuveline, P., J. Timberlake, and F. Furstenberg. 2003. "Shifting Child Rearing to Single Mothers: Results from 17 Western Nations." *Population and Development Review* 29: 47-71.
- Jacobs, J. and R. King. 2002. "Age and College Completion: A Life-History Analysis of Women Aged 15-44." *Sociology of Education* 75(3): 211-230.
- Jencks, C., S. Bartlett, M. Corcoran, J. Crouse, D. Eaglefield, G. Jackson, K. McClelland, P. Meuser, M. Olneck, J. Scharwitz, S. Ward, and Jill Williams. 1979. *Who Gets Ahead? The Determinants of Economic Success in America*. New York: Basic.
- Kane, T. and C. Rouse. 1995. "Labor-Market Returns to Two- and Four-Year College." *American Economic Review* 85: 600-614.
- Karlson, K. B., A. Holm, and Richard Breen. 2012. "Comparing Regression Coefficients Between Same-Sample Nested Models Using Logit and Probit: A New Method." *Sociological Methodology* 42(1): 286-313.
- Kaufman, Philip, Martha Naomi Alt, and Christopher Chapman. 2001. "Dropout Rates in the United States: 2000." Washington, D.C.: U.S. Government Printing Office.
- Kerckhoff, A. 1976. "The Status Attainment Process: Socialization or Allocation?" *Social Forces* 55: 368-381.
- Kerckhoff, A. 1993. *Diverging Pathways: Social Structure and Career Deflections*. New York: Cambridge University Press.
- Kerckhoff, A. and L. Bell. 1998. "Hidden Capital: Vocational Credentials and Attainment in the United States." *Sociology of Education* 71:152-174.
- Kerckhoff, A. 2001. "Education and Social Stratification Processes in Comparative Perspective." *Sociology of Education* 74 Extra Issue: 3-18.
- Kerckhoff, A. and E. Glennie. 1999. "The Matthew Effect in American Education." *Research in Sociology of Education and Socialization* 12: 35-66.
- Kowaleski-Jones, L. and Rachel Dunifon. 2006. "Family Structure and Community Context: Evaluating Influences on Adolescent Outcomes." *Youth and Society* 38(1): 110-130.

- Kuo, H. and R. Hauser. 1995. "Trends in Family Effects on the Education of Black and White Brothers." *Sociology of Education* 68:136-160.
- Land, K. and S.T. Russell. 1996. "Wealth Accumulation Across the Adult Life Course: Stability and change in sociodemographic covariate structures of net worth data in the Survey of Income and Program Participation, 1984-1991." *Social Science Research* 25: 423-462.
- Lang, K. and Jay Zagorsky. 1997. "Does Growing Up with a Parent Absent Really Hurt?" Working Paper, Columbus OH: The Ohio State University, Center for Human Resource Research, October 1997
- Lareau, A. 2003. *Unequal Childhoods: Class, Race-ethnicity, and Family Life*. Berkeley: UC Press.
- Lareau, A. and D. Conley. 2008. *Social Class: How Does it Work?* Russell Sage Foundation.
- Lawlor, D., M. Shaw, and S. Johns. 2001. "Teenage pregnancy is not a public health problem." *British Medical Journal* 323(7326): 1428-1429
- Loh, E. 1996. "Changes in Family Structure, Attained Schooling, and Adult Poverty Status" *Social Science Quarterly* 77,1: 145-158
- Long, J. Scott. 1997. *Regression Models for Categorical and Limited Dependent Variables*. Thousand Oaks, CA: Sage.
- Long, J. S. and Jeremy Freese. 1997. *Regression Models for Categorical and Limited Dependent Variables*. Advanced Quantitative Techniques in the Social Sciences 7. Thousand Oaks, CA: Sage Publications.
- Lorence, J., A. Dworkin, L. Toenjes, and A. Hill. 2002. "Grade Retention and Social Promotion in Texas, 1994-1999: Academic Achievement among Elementary School Students." In *Brookings Papers on Educational Policy: 2002*, edited by D. Ravitch, pp. 13-52, Washington, DC: The Brookings Institution.
- Loveless, T. 1999. *The Tracking Wars: state Reform Meets School Policy*. Washington, DC: Brookings Institution Press.
- Luke, Douglas. 2004. *Multilevel Modeling*. Thousand Oaks: Sage.
- Luker, K. 1996. *Dubious Conceptions: The politics of Teenage Pregnancy*. Harvard University Press, Cambridge, MA.

- MacKinnon, D. A. Fairchild, and Matthew Fritz. 2007. "Mediation Analysis." *Annual Review of Psychology* 58: 593-603.
- Macunovich, D. 1996. "A Review of Recent Developments in the Economics of Fertility." Pp. 91-158 in *Household and Family Economics*. P. Menchik (ed.). Dordrech: Kluwer Academic.
- Maralani, V. 2011. "From GED to College: Age Trajectories of Non-Traditional Educational Paths." *American Educational Research Journal* 20(10): 1-32.
- Mare, R.D. 1980. "Social Background and School Continuation Decisions." *Journal of the American Statistical Association* 75(2): 295-305.
- Mare, R.D. 1981. "Change and Stability in Educational Stratification." *American Sociological Review* 46: 72-87.
- Martin, J.A., B.E. Hamilton, and S.J. Ventura. 2015. *Births: Final Data for 2013*. Hyattsville, MD: National Center for Health Statistics
- Mayer, K. and N. Tuma. 1990. *Event History Analysis in Life Course Research*. Madison, WI: University of Wisconsin Press.
- McLanahan, S. 1985. "Family Structure and the Reproduction of Poverty." *American Journal of Sociology* 90(4): 873-901.
- McLanahan, S. 1997. "Parent Absence or Poverty: Which Matters More?" pp. 35-48 in *Consequences of Growing Up Poor*, Duncan, G. and J. Brooks-Gunn (eds.) New York: Russell Sage Foundation.
- McLanahan, S. and Gary Sandefur. 1994. *Growing Up With a Single Parent: What Helps and What Hurts*. Cambridge, MA: Harvard Press.
- Menard, S. 2002. *Applied Logistic Regression Analysis*. Quantitative Applications in the Social Sciences Series. Thousand Oaks: Sage.
- Moore, K., D. Myers, D. Morrison, and B. Edmonston. 1993. "Age at First Childbirth and Later Poverty." *Journal of Research on Adolescence* 3,4: 393-422.
- Moore, K. and L. Waite. 1977. "Early Childbearing and Educational Attainment." *Family Planning Perspectives* 9(5): 220-225.

- Mouw, T. 2005. "Sequences of Early Adult Transition: A Look at Variability and Consequences." Pp. 256-291. *On the Frontier of Adulthood: Theory, Research, and Public Policy*. Frank Furstensburg, Jr., Ruben Rumbaut and Richard Settersten, Jr., editors. University of Chicago Press.
- Mueller, D. P., and P.W. Cooper. 1986. "Children of single parent families: How they fare as young adults." *Family Relations* 35: 169-176.
- Mulder, C.H. 2006. "Homeownership and Family Formation." *Journal of Housing and the Built Environment* 21:281-298.
- Murnane, R., J. Willett, J. and F. Levy. 1995. "The Growing Importance of Cognitive Skills in Wage Determination." *Review of Economics and Statistics* 77:251-266.
- Murnane, R., K. Boudett, and J. Willett. 1999. "Do Male Dropouts Benefit from Obtaining a GET, Postsecondary Education, and Training?" *Evaluation Review* 23: 475-502.
- National Center for Education Statistics (NCES). 2005. "Waiting to Attend College: Undergraduates Who Delay Their Postsecondary Enrollment." NCES 2005-152. Washington, D.C.: U.S. Department of Education.
- National Longitudinal Survey of Youth. 1983. *NLSY79: Technical Sampling Report*. (M. Frankel, H. McWilliams, and B. Spencer, authors). Columbus, OH: NLS User Services.
- National Research Council. 1999. In Heubert, J. and R.M. Hauser (eds.), *High Stakes: Testing for Tracking, Promotion, and Graduation*. Washington, DC: National Academy Press.
- Natriello, G. 1994. "Coming Together and Breaking Apart: Unifying and Differentiating Processes in Schools and Classrooms." *Research in Sociology of Education and Socialization* 10: 111-145.
- Nemes, S., J. Jonasson, A. Genell, and Gunnar Steineck. 2009. "Bias in Odds Ratios by Logistic Regression Modeling and Sample Size." *BMC Medical Research Methodology* 9:56.
- Oliver, M. and Thomas Shapiro. 2006. *Black Wealth/White Wealth*. New York: Routledge.
- Pallas, A. 1993. "Schooling in the Course of Human Lives: The Social Context of Education and the Transition to Adulthood in Industrial Society." *Review of Educational Research* 63(4): 409-447.
- Pallas, A. 2004. "Educational Transitions, Trajectories, and Pathways." Pp. 165-184 in Mortimer, J. and M. Shanahan (eds.) *Handbook of the Life Course*. Springer.

- Painter, M. and Kevin Shafer. 2007. "Children, Family Size Change, and Household Wealth Trajectories." Presented: New York, NY, Population Association of America Annual Meetings, March 29-31.
- Peters, E. and N. Mullis. 1997. "The Role of Family Income and Sources of Income in Adolescent Achievement." pp. 340–381 in *Consequences of Growing Up Poor*, Duncan, G. and J. Brooks-Gunn (eds.) New York: Russell Sage Foundation.
- Plunkett, S. and M. Bamaca-Gomez. 2003. "The Relationship Between Parenting, Acculturation, and Adolescent Academics in Mexican-origin Immigrant Families in Los Angeles." *Hispanic Journal of Behavioral Sciences* 25: 222-239.
- Preacher, K. and Andrew Hayes. 2008. "Asymptotic and Resampling Strategies for Assessing and Comparing Indirect Effects in Multiple Mediator Models." *Behavior Research Methods* 40(3): 879-891.
- O'Connor, James. 1973. *The Fiscal Crisis of the State*. New York: St. Martin's Press.
- O'Rand, A. 1996. "The Precious and the Precocious: Understanding Cumulative Advantage and Cumulative Disadvantage Over the Life Course." *Gerontologist* 36:230-238.
- O'Rand, A. 2002. "Cumulative Advantage Theory in Life Course Research." *Annual Review of Gerontology and Geriatrics* 22:14-20.
- Powell, B. and L.C. Steelman. 1990. "Beyond sibship size: sibling density, sex composition, and educational outcomes." *Social Forces* 69: 181–206.
- Powell, B. and L.C. Steelman. 1993. "The educational benefits of being spaced out: sibship density and educational progress." *American Sociological Review* 58: 367–381.
- Qian, Z. and S. Blair. 1999. "Racial/ethnic Differences in Educational Aspirations of High School Seniors." *Sociological Perspectives* 42: 605-625.
- Roksa, J. and M. Velez. 2012. "A Late Start: Delayed Entry, Life Course Transitions and Bachelor's Degree Completion." *Social Forces* 90: 769-794.
- Sandefur, G., S. McLanahan, and R. Wojtkiewicz. 1989. "Race-ethnicity and Ethnicity, Family Structure, and High School Graduation." Institute for Research on Poverty Discussion Paper 893-89. Madison: University of Wisconsin.
- Shaff, K., N. Wolfinger, L. Kowaleski-Jones, and Ken R. Smith. 2008. "Family Structure Transitions and Child Achievement." *Sociological Spectrum* 28: 681-704.

- Shavit, Y. & H. P. Blossfeld. 1993. *Persistent Inequality: Changing Educational Attainment in Thirteen Countries*. Boulder, CO: Westview Press.
- Sewell, W., and R. Hauser. 1980. "The Wisconsin Longitudinal Study of Social and Psychological Factors in Aspirations and Achievements." In A.C. Kerckhoff (ed.), *Research in Sociology of Education and Socialization* 1:59-99. Greenwich, CT: JAI Press.
- Singer, J. and J. Willett. 2003. *Applied Longitudinal Data Analysis: Modeling Change and Event Occurrence*. Cambridge: Oxford Press.
- Teachman, J., K., Paasch, R. Day, and K. Carver. 1997. "Poverty During Adolescence and Subsequent Educational Attainment." In *Consequences of Growing Up Poor* G.J. Duncan and J. Brooks-Gunn (eds.). New York, NY: Russell Sage Foundation: 382-418.
- Upchurch, D. and J. McCarthy. 1990. "The Timing of a First Birth and High School Completion." *American Sociological Review* 55(2): 224-234.
- van der Gaag, J. and E. Smolensky. 1982. "True Household Equivalence Scales and Characteristics of the Poor in the United States." *Review of Income and Wealth* 28(1): 17-28.
- Veum, J. R., and A.B. Weiss. 1993. "Education and the work histories of young adults." *Monthly Labor Review* 116(4):11-20.
- Wagmiller, R., E. Gershoff, P. Veliz, and Margaret Clemens. 2010. "Does Children's Academic Achievement Improve When Mothers Marry?" *Sociology of Education* 83(3): 201-226.
- Warren, J., J. Sheridan, and R. Hauser. 2002. "Occupational Stratification Across the Life Course: Evidence from the Wisconsin Longitudinal Study." *American Sociological Review* 67: 4323-455.
- Wao, H. 2010. "Time to the Doctorate: Multilevel Discrete-Time Hazard Analysis." *Educational Assessment, Evaluation, and Accountability* 22(3): 227-247.
- Weber, Max. 1923 (1961). *General Economic History*. Translated by Frank H. Knight. New York City: Collier-Macmillan.
- Willett, John B., and Judith D. Singer. 1991. "From Whether to When: New Methods for Studying Student Dropout and Teacher Attrition." *Review of Educational Research* 61:407-450.



- Williams, R. 2009. "Using Heterogeneous Choice Models to Compare Logit and Probit Coefficients across Groups." *Sociological Methods and Research* 37(4):531–59.
- Wolfe, B., R. Haveman, D. Ginther, and C. Bum. "The 'Window' Problem in Studies of Children's Attainments: A Methodological Exploration." *Journal of the American Statistical Association* 91.
- Wu, Lawrence L. 2003. "Event History Models for Life Course Analysis." Pp. 477-502. *Handbook of the Life Course*. Jeylan T. Mortimer and Michael J. Shanahan, editors. Springer.
- Yamokoski, A. and Lisa Keister. 2006. "Wealth of Single Women: Marital Status and Parenthood in the Asset Accumulation of Young Baby Boomers in the United States." *Feminist Economics* 12,1-2 (January-April 2006):167-194. A Special Issue on Women and Wealth
- Zagorski, J. 2005. "Marriage and Divorce's Impact on Wealth" *Journal of Sociology* 41,4: 406-424.
- Zagorski, J. 2010. Personal email communication, September 1.

## APPENDIX A

### **Nature of the Data, Definitions for Variables, and a Note on “Race-ethnicity” in the NLSY**

#### *Nature of the Data*

Although “observed characteristics” are coarse proxies for real phenomena, survey data do provide many well-tested and meaningful markers that must be used in place of better information about people’s lives. (For example, observable features of family conditions such as parental wealth and annual family income cannot describe any particular family’s economic conditions with perfect accuracy, but they do approximate such conditions offer a glimpse into how family conditions can vary.) The National Longitudinal Survey of Youth 1979 cohort (NLSY) is a complex panel survey that originally included a nationally representative sample of 12,686 men and women aged 14 to 21 years on December 31, 1978. NLSY is a multistage, stratified, clustered probability sample meant to represent the entire population of youth residing in the United States in 1979. On the heels of earlier (late 1960s) attempts at collecting longitudinal data, the Bureau of the Census and the Bureau of Labor Statistics initiated this study initially to enhance our knowledge of Americans’ labor market experiences. The survey contains many years of information covering educational attainment, employment history, training investments, income and assets, welfare receipt, child-care costs, insurance coverage, health conditions, workplace injuries, alcohol and substance abuse, sexual activity, and marital and fertility histories.

The sample respondents examined in this dissertation were re-interviewed annually from 1979 through 1994 and then biennially until 2004 (as indicated in the sections above, I used data for the time period 1979 to 2004). NLSY has important advantages for an event history analysis

because data about the timing of educational attainment and many of its potential predictors were collected contemporaneously at each survey wave, which gives rich detail and enhances accuracy. I restricted the sample to respondents who remained in the survey over a period of 25 years, from 1979 to 2004.

The statistics package STATA/IC version 12.1 was used for all aspects of data management as well as the computations, analyses, and generating tables and graphs. In brief, after creating a “person-period” file for the data and generating dummy age indicators for a general (nonparametric) time specification on the “college completion” variable, a general baseline model was estimated in which several background variables (age, race-ethnicity, sex) predicted college completion. Then, using life course theory and existing scholarship as a guide, full models were specified by successively adding explanatory variables such as maternal education, family poverty, teenage parenthood, number of siblings, and so forth.

### *Definitions for Variables*

The list in Table C1 below provides succinct definitions for variables used in my preliminary investigations as well as in all final models included in the modeling and analysis sections above. Immediately below the list of variable definitions I also included a brief discussion and rationale for how I created the “race-ethnicity” variable based on two of the existing variables in NLSY.

**Table A1**  
**Existing/Created Variables for Initial Exploratory Analyses and Final Models**

<b>Variable</b>	<b>Definition</b>
<i>id</i>	Unique identifier (numeric) for each respondent
<i>gender</i>	Gender of respondent (self-reported)
<i>race-ethnicity*</i>	Race-ethnicity of respondent (white, black, Hispanic)
<i>dobmonth</i>	month of respondent's birth
<i>dobyyear</i>	calendar year of respondent's birth
<i>age*</i>	Age in years during listed year (calculated from given birthdate)
<i>everchildren1979</i>	Indicator of whether respondent had ever had a child as of 1979
<i>teenparent*</i>	Indicator of whether respondent became a parent prior to age 20
<i>agefirstbirth*</i>	Age at which respondent experienced the birth of first child
<i>haschild*</i>	time-varying indicator of whether respondent is a parent
<i>livedparents*</i>	Indicator of whether respondent lived with both biological parents from birth to age 18
<i>wealth*</i>	Wealth (net worth, in constant dollars) of respondent's family-of-origin (during listed years 1979 to 1987)
<i>anypov*</i>	Indicator of whether family-of-origin was below the poverty line during at least one year during respondent's adolescence
<i>eduattain</i>	R's years of completed education during listed year
<i>eduage25*</i>	Respondent's years of completed education by age 25
<i>agehsgrad*</i>	Age at which respondent completed high school or GED
<i>timetocollgrad*</i>	Observed time (in years) until respondent completed college (or until censoring)
<i>collgrad*</i>	Categorical recode of "timetocollgrad"
<i>fampoverty*</i>	Family poverty status (past calendar year, based on U.S. threshold for given year)
<i>eduattain25</i>	Highest level of education by the time respondent reached age 25 (as indicated by "degree name," bachelors degree, associates degree, high school diploma, etc.)
<i>eduattain</i>	level of education reached by respondent (as indicated by "degree name," bachelors degree, associates degree, high school diploma, etc.)
<i>marital</i>	Respondent's marital status during listed year
<i>motheredu</i>	Respondent's mother's years of education in listed year
<i>fatheredu</i>	Respondent's father's years of education in listed year
<i>motheryears*</i>	Indicator of how many years respondent's mother was present in the household (until respondent turned age 19)
<i>fatheryears*</i>	Indicator of how many years respondent's father was present in the household (until respondent turned age 19)
<i>siblings1979</i>	number of siblings respondent had in the year 1979
<i>siblings</i>	Number of siblings respondent had (during however many years she was living with the family-of-origin)
<i>sibedu</i>	Highest grade completed by respondent's oldest sibling
<i>oldersibs</i>	Number of siblings older than respondent
<i>famincome*</i>	Family income in constant dollars (past calendar year)
<i>famsize</i>	Respondent's family size (number of people) during listed year
<i>intactfam*</i>	Indicator of whether respondent lived with both parents to age 18

**Table A1 (continued)**  
**Existing/Created Variables for Initial Exploratory Analyses and Final Models**

Variable	Definition
<i>famstructchange*</i>	Change in parental presence, indicating several possibilities including “change in parental presence due to death of father,” or “loss of father due to divorce,” and so forth.

\* Indicates a created variable (one not originally in NLSY but calculated from existing variables). All details, data files, and Stata code are available upon request.

**\*\*Note on “Race-ethnicity” in the NLSY**

The variable for racial identification used in this dissertation was not based on the initial NLSY determination and does *not* presume to indicate an individual’s personal racial or ethnic identity. (As with so many social phenomena, identifying a person’s race-ethnicity can be a very tricky business depending on how one conceptualizes “race-ethnicity.”) Rather, I created the variable *race-ethnicity* by sorting through two different race-ethnicity variables existing in NLSY, namely “race-ethnicity” (how the interviewer categorized the respondent’s phenotypic characteristics) and “race-ethnicity primary” (how the respondent personally identified his/her racial identity). In cases where there was a mismatch between the two designations, I allowed the interviewer’s judgment to supercede the respondent’s stated self-identity. The assumption here is that an interviewer’s initial visual determination of someone’s so-called race-ethnicity is plausibly similar to how others in social life (secondary contacts, especially) may see or “read” that person in racial terms. In terms of life chances, personal racial feeling or identity is less consequential than the social significance of typically “black” phenotypic characteristics. Since my analyses do not reckon with the issue of personal racial identity, I determined that black social identity (how others read a person’s apparent “race-ethnicity” even if a respondent self-identifies otherwise) should trump other non-white classifications in cases where the interviewer reported a respondent’s race-ethnicity as black.

## APPENDIX B

### **Note on the Preliminary Inquiry into Family Transitions and Educational Attainment**

Family structure and family background are not identical, the former being a key dimension of the latter. Initially I focused on *parental situation* (an aspect of “family structure”) that existed in the household during an individual’s formative years from birth to age 18, and I did so by dividing parental situation into various types described below. (Due to sample size limitations, I excluded families headed by single fathers.) Two preliminary research questions were:

- 1) Does “change in family structure” *per se* affect individuals odds of completing college?
- 2) Do certain *types* of change in family structure affect odds of educational completion more than other types of change?
- 3) Does the effect of certain types of family structure change depend on *when* they occur in an individual’s life?

The motivation for this initial line of inquiry was that some researchers have pointed to specific types of family change (particularly divorce resulting in the loss of a father figure) as having presumably negative effects on children’s educational attainment. The argument is that we should expect this negative influence to be most acute when divorce occurs during children’s teenage years (Loh 1995). As well, other studies of educational attainment emphasize socialization dynamics and demonstrate the more generalized stresses associated with changes in family structure *per se*. My interest was in the possible negative effect such stresses may have depending on when in a person’s life a transition in parental situation occurs; compared to earlier-life changes in family structure at, say, age 5 or 6, transitions that occur later (say, during

an individual's adolescence) may be more detrimental to her likelihood of completing high school or college because the disruption caused by earlier changes may fade before she is even at risk of quitting high school or enrolling in college (McLanahan 1985). Others have argued that the age-varying need for and availability of economic resources shapes the way family structure effects may vary with offspring age (Cameron and Heckmand 1991; Loh 1995; Shaff et al. 2008).

In any case, first I summarized the range of possible family structure transitions and the time periods in respondents' lives when family change occurred. Next, I examined the associations between family structure changes and educational attainment. Initially I limited the definition of family change simply to whether there was a change in family structure at all during childhood. The NLSY data offer greater detail than family change *per se*, so I was able to observe some reasons for family change (parental death, divorce or separation, remarriage, etc). I limited the definition of family change to whether an individual experienced a change in family structure during childhood (dummy: separated from one or both parents during the first 18 years of life. Comparison group: lived with both birth parents until age 18). For greater detail, I also categorized some potential reasons for family change (death of father/mother, loss of father/mother due to divorce or separation, never knew father, other).

When the instability hypothesis holds forth in studies of family structure, substantial policy implications tend to be assumed. The literature on single-parenthood provides one of the more glaring illustrations of what the instability hypothesis implies: A child born to a single parent might be as well off, or perhaps even better off, if the parent does not subsequently cohabit or marry. According to Fomby and Cherlin (2007, 181), this would require social scientists to significantly reconsider the previously reported "effects" of family structure on

various child outcomes. The sociology of education and socialization literatures offer ways to theorize the links between family structure and educational attainment (e.g., Bryant 1990; McLanahan and Bumpass 1988; Sandefur et al. 1989; Garasky 1995; Fomby and Cherlin 2007). Based on this literature (McLanahan 1985; Fomby and Cherlin 2007), one expectation is that the stresses associated with a change in family structure *per se* should negatively affect timely educational attainment. However, in analyses not shown, I modeled the timing of family structure transitions (loss of parent due to death of father or mother during childhood or adolescence, loss of parent due to divorce during childhood or adolescence, and so forth as described above) and found minimal to no effects of such family structure transitions on individuals' risks of college completion by the time they were 25 years old ("eduage25" as displayed by Table C1 in Appendix C).<sup>51</sup>

---

<sup>51</sup> McLanahan (1985) showed that the stress of a family structure transition is temporary and often diminishes over time. More recently however, Fomby and Cherlin (2007) used NLSY data to show that family change *per se*—regardless of the type of transition or when it occurs—has a persistent and generally negative effect on key child cognitive and behavioral outcomes; by extension, their findings suggest that family change may indirectly affect educational attainment.



## APPENDIX C

### **Note on Initial Sample Statistics, Multicollinearity, and Decisions About Variables**

Initially the sample statistics showed mother's education and age at first pregnancy as having the strongest positive correlation with educational attainment (see Table A1). Considering the prevalence of endogamy in the United States, and especially marital homogamy by socioeconomic status, it is unsurprising to observe the strong correlation between mother's and father's educational attainment. (This was but one reason for my decision to drop father's education from the main analyses.) There were also negative correlations between number of siblings and educational attainment, and between number of siblings and age at first pregnancy. This is unsurprising as well, since other studies have found connections between family size and within-family competition for scarce emotional and material resources, cultural capital, parental time, and so forth (e.g., see Garasky 1995). The other correlations among just these initial predictors were relatively small, indicating a low degree of multicollinearity and that the predictors were efficient and unbiased predictors of educational attainment. It was with these conditions met that I conducted several exploratory regression analyses.

**Table C1**  
**Initial Descriptive Statistics and an Exploratory Regression Analysis for “Years of Completed Education by Age 25”**

White respondents			Black respondents		
<u>Variable</u>	<u>mean</u>	<u>s.d.</u>	<u>Variable</u>	<u>mean</u>	<u>s.d.</u>
AgeFirstPreg	22.06	5.23	AgeFirstPreg	19.97	4.94
EduAge25	13.13	2.22	EduAge25	12.50	1.89
motheryears	17.87	3.47	motheryears	17.40	4.31
fatheryears	16.01	5.53	fatheryears	12.03	7.94
siblings	3.30	2.15	siblings	4.77	3.09

Hispanic respondents		
<u>Variable</u>	<u>mean</u>	<u>s.d.</u>
AgeFirstPreg	21.11	5.03
EduAge25	11.83	2.75
motheryears	17.81	3.62
fatheryears	14.78	6.67
siblings	4.55	3.05

Exploratory Regression Analysis for “Years of Completed Education by Age 25”									
	<u>Edu~25</u>	<u>AgePrg</u>	<u>Mothedu</u>	<u>FathEdu</u>	<u>sibEdu</u>	<u>mothyrs</u>	<u>fathyrs</u>	<u>siblings</u>	
<b>EduAge25</b>	1.0000								
<b>AgeFirstPreg</b>	0.4878	1.0000							
<b>MotherEdu</b>	0.4489	0.2510	1.0000						
<b>FatherEdu</b>	0.4483	0.2094	0.6441	1.0000					
<b>sibEdu</b>	0.2292	0.1332	0.2978	0.3532	1.0000				
<b>motheryears</b>	0.1278	0.1163	0.0242	0.0430	0.0394	1.0000			
<b>fatheryears</b>	0.1791	0.2098	0.0938	0.0449	0.0631	0.3744	1.0000		
<b>siblings</b>	-0.2402	-0.1638	-0.3119	-0.3561	-0.1424	-0.0062	-0.0234	1.0000	

(all correlations statistically significant at  $p < .05$ )

Using educational attainment as the response variable (measured in years of completed education at a respondent’s 25th birthday), I ran regressions to determine which family background variables might best predict educational attainment among respondents in the sample. These initial independent variables were race-ethnicity, number of years the mother had spent in the household while respondent was growing up, number of years the father had spent in the household, “home literacy environment” (presence of magazines, newspapers, and library

card), age at first pregnancy, mother's and father's educational attainment, and educational attainment of the respondent's oldest sibling. (Considered altogether, these predictors explained 39% of the variation in respondents' years of education by age 25. The relative power of these variables is notable mainly because my analyses had not yet included other logically and theoretically important predictors such as family poverty during adolescence, parental wealth, number of siblings in a respondent's household, and so forth.) Black and Hispanic race-ethnicity/ethnicity both were negatively correlated with years of education by age 25 ("white" as dummy); the typical Black respondent had over half a year less education at age 25 compared to a typical white respondent, and a typical Hispanic respondent had over a year less education compared to a typical white. The duration of each biological parent's presence in the respondent's household, as measured by the *motheryears* and *fatheryears* variables, was very weakly related to educational attainment, and the second column from left shows that controlling for "years a parent was present in the household" (*motheryears* and *fatheryears*) only slightly attenuated the relationship between racial minority status and educational attainment, thus these "parental years" predictors were dropped from the final models that appeared in the main analyses above.

Models 3 and 4 in Table A2 below show that both home literacy environment (HL: the grouped effect of three variables indicating "presence of library card, newspapers, or magazines in an individual's home") and age at first pregnancy (AFP) altered the association between race-ethnicity and educational attainment. Specifically, HL and AFP conditioned the respective effects of Black and Hispanic race-ethnicity on educational attainment. When the set of three HL variables were included (Model 3), the effect of Black race-ethnicity on educational attainment reversed direction from negative to weakly positive (the coefficient for educational attainment

changed from -0.552 to 0.089). In that same model, the effect of Hispanic race-ethnicity on educational attainment changed similarly with the introduction of HL, but it did not change *direction* until the block of two parental education variables (*motheredu* and *fatheredu*) were introduced (Model 5). This finding, along with a desire for parsimony, is why I dropped the HL variables from final models that appeared in the main analyses above.

Looking to Model 6 in Table A2 below, both Hispanic ethnicity and black race-ethnicity appeared to positively affect educational attainment (with the 0.528 coefficient, black race-ethnicity seemed to positively affect educational attainment by more than half a year). However, sociological theories of educational stratification suggest that this is highly implausible; indeed, upon further inspection, it seemed that the Hispanic and black variables had *indirect* effects on educational attainment. Specifically and when comparing the successive models not shown here, Black and Hispanic race-ethnicity-ethnicity positively affected educational attainment by age 25 only indirectly through the age at first pregnancy (AFP) variable. It was for this reason and several others (discussed in the sections above) that I included teenage childbearing in the analyses.

**Table C2**

**Regression Analysis for “Years of Completed Education by Age 25” (with some predictors added sequentially in blocks)**

<b>VARIABLE</b>	<b><u>Model 1</u></b>	<b><u>Model 2</u></b>	<b><u>Model 3</u></b>	<b><u>Model 4</u></b>	<b><u>Model 5</u></b>	<b><u>Model 6</u></b>
Black	-0.552*** (0.161)	-0.447*** (0.162)	0.0890 (0.161)	0.374** (0.147)	0.519*** (0.139)	0.528*** (0.138)
Hispanic	-1.305*** (0.203)	-1.318*** (0.202)	-0.571*** (0.202)	-0.455** (0.184)	0.360* (0.185)	0.281 (0.185)
motheryears		0.0514** (0.0203)	0.0383** (0.0194)	0.0133 (0.0177)	0.0177 (0.0165)	0.306*** (0.0820)
fatheryears		0.0350*** (0.0118)	0.0208* (0.0113)	0.00407 (0.0103)	0.00555 (0.00967)	-0.0620** (0.0302)
magazines			0.965*** (0.123)	0.690*** (0.113)	0.455*** (0.108)	0.446*** (0.107)
newspapers			0.775*** (0.158)	0.683*** (0.143)	0.255* (0.138)	0.260* (0.137)
library			0.442*** (0.133)	0.328*** (0.121)	-0.0124 (0.116)	-0.0210 (0.115)
AgeFirstPreg				0.155*** (0.00895)	0.129*** (0.00856)	0.127*** (0.00854)
FatherEdu					0.104*** (0.0168)	0.00392 (0.0478)
MotherEdu					0.138*** (0.0220)	0.581*** (0.126)

Standard errors in parentheses, with \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

## APPENDIX D

### General Note on Wealth and Discussion of Data Limitations

Examining both income and wealth provides a more complete picture of economic well-being.

The NLSY data represent a unique source of wealth information because wealth information was gathered over many years in people's lives and especially because they provide good measures of wealth for low-income and minority households. [Note: For precision, much of the discussion below is drawn from the *NLSY79 User's Guide* (Bureau of Labor Statistics 2001, 108-110).]

Since 1979 the NLSY respondents have been asked various questions about assets, debts, when saving begins, how savings habits are formed, and how persistent savings habits are. But the NLSY's wealth data were somewhat problematic for my interest in wealth accumulation over the life course (for substantive reasons I did not trace wealth during respondents' earliest or later years) due to low response rates in certain years and inconsistency regarding the questions asked.

#### *Limited Years of Wealth Data*

As mentioned in the Section 4 discussion, there are reasons for including only a limited number of years for the wealth variables in the analyses. First, each of the first four surveys (1979 to 1982) contained questions asking if the respondent or their spouse had any money set aside for savings, owned a vehicle, or owned their own home. Unfortunately however, respondents were never asked how much savings were held, or the estimated value or number of vehicles, or the value of their home, or the value of their mortgage. Second, during these same survey years, respondents were only asked questions on assets if they met at least one of the following five criteria:

- 1) were 18 years or older
- 2) had a child
- 3) were enrolled in college
- 4) were married, or
- 5) were living outside their parents' home.

This selection process eliminated many respondents from being eligible to answer the wealth questions during those early years. Indeed, early NLSY data show that very few individuals answered the questions before they turned 18 years old.

Beginning in 1985 the wealth data are more useful, as respondents have been administered a larger and more detailed wealth section (approximately 20 questions) than before, so these data provide estimates of how much individuals' homes, cash savings, stock and bond portfolios, estates, businesses, and automobiles were worth. As well, respondents have been asked to estimate how much mortgage debt, property debt, and other debt they had accumulated. Starting in 2004 the wealth module (set of variables) in NLSY 1979 became considerably more complex, thus I limited the analysis period to 1985 – 2004 (excluding 1991 and 1995). These years provide a good look at each respondent's net worth in their post-adolescent but pre-retirement lives. The years 1991 and 1995 were periods during which two breaks occurred in the wealth time series. As indicated in Zagorski's diagnostic paper on NLSY wealth data (1999, 137), budgetary restrictions in the survey's 13th round "resulted in all wealth questions being eliminated in 1991 for one round of questioning. There are no data for 1995 since the NLSY switched from interviewing respondents every year to every other year, beginning in 1994, to lower the survey's cost and reduce respondent burden."

Income most often is depicted as a "flow" of payment (units usually are dollars/time). Wealth can be conceptualized and even measured similarly, as a potential income flow. One way

to establish basic comparability between wealth and income is to measure wealth as a “capitalized income stream.” With statistical code provided by the Center for Human Resource Research, I converted individuals’ net worth/assets into a yearly payment stream. This was done using an amortization calculation, with a 7% fixed interest rate compounded each year, which created a longitudinal series of net asset variables for the years relevant for my analyses. Wealth is conceptualized as an individual’s family’s *liquid financial assets* (stocks, bonds, trusts, retirement saving, cash savings), *illiquid assets* such as automobiles or business/farm equity, and *residential assets* (home value minus value of the mortgage and any property debt). I also imputed for missing values, summed the various components of people’s net assets, and adjusted for inflation using CPI for given years.

#### *Top Coding for Wealth and Income*

Since the NLSY is publically available to the research and public policy communities, extensive measures are taken to protect respondents from being identified. As is typical with such datasets, the Bureau of Labor Statistics has used various algorithms over the years to “top code” any unusually high values (this holds true for all wealth and income variables) to ensure confidentiality.