THREE ESSAYS ON THE MICRO BASIS OF SOCIOECONOMIC INEQUALITY:
The Role of Cognitive and Noncognitive Skills

Dohoon Lee

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Sociology.

Chapel Hill
2008

Approved by:

Kathleen Mullan Harris

Advisor: Kathleen Mullan Harris

Guang Guo

Reader: Guang Guo

Reader: Philip Cohen

Ted Mouw

Reader: Ted Mouw

Reader: Krista Perreira

©2008
Dohoon Lee
ALL RIGHTS RESERVED

ABSTRACT

DOHOON LEE: Three Essays on the Micro Basis of Socioeconomic Inequality:
The Role of Cognitive and Noncognitive skills
(Under the direction of Kathleen Mullan Harris)

This dissertation explores the effect of cognitive and noncognitive skills on three socioeconomic outcomes: wage differentials, individual patterns of educational assortative mating, and the socioeconomic consequences of teen motherhood. Although research has been keen on identifying early predictors of socioeconomic attainment, a systematic view of the linkages between individuals' own attributes formed in childhood and adolescence and subsequent outcomes has yet to come. In this project, I seek to fill this gap by identifying cognitive and noncognitive skills as a micro basis of socioeconomic inequality. Correlated but distinct from cognitive skills, noncognitive skills are conceptualized as enduring dispositions that represent a form of cultural capital. Because both types of skills are highly dependent on socioeconomic background, I hypothesized that individual-level skill differences function as a key channel through which intergenerational mobility is associated with various forms of socioeconomic inequality.

This study begins by examining the impact of both cognitive and noncognitive skills on between- and within-education group wage inequality, using data from the National Longitudinal Survey of Youth 1979 (NLSY79) and quantile regressions. While the economic return to education has been proposed as a parsimonious explanation for rising wage inequality and its currently high level, this account focuses mainly on between-education group wage inequality and the demand for cognitive skills. Results from my analysis shows that 1) while the economic return to education is the robust explanation for wage inequality,

both cognitive and noncognitive skills contribute significantly to reducing the college wage premium and wage dispersions within college graduates; 2) noncogntive skills play a more pronounced role in wage inequality among college graduates; and 3) the wage effect of both skills as an early predictor of earnings strengthens as workers reach their prime ages in the labor market. These findings suggest that the family may be an important institutional actor responsible for wage inequality.

In the subsequent chapter, I use data from the National Longitudinal Study of Adolescent Health (Add Health) and the NLSY79 to investigate the role of cognitive and noncognitive skills in individual patterns of educational assortative mating in adolescence and adulthood. This paper argues that to the extent that skill differences are associated with education-based mate selection, intergenerational mobility operates at an intimate level of the mate selection process. Multinomial logistic regression results show that cognitive and noncognitive skills are positively associated with the probability of transitioning into marrying college graduates and, to a lesser degree, with that of dating college-bound partners. I also find a gender difference in the role of skill differences: noncognitive skills play a more salient role in education-based mate sorting for women, whereas it is cognitive skills that primarily do so for men, indicating a normative attitude toward mate selection that regards “smartness” as more attractive to women than to men. These findings imply that the intergenerational transmission of familial resources affects children’s mate selection by not simply investing in their educational attainment but also strengthening their cognitive and noncognitive skills.

Finally, I reevaluate the socioeconomic effect of teenage childbearing. Despite a 30-year debate about the consequences of adolescent fertility, finding its “true” effect still has been elusive. This concern stems from 1) theoretical considerations of early motherhood as a harmful event and/or its higher likelihood among disadvantaged young women and 2) methodological challenges against selection bias. Alternative models have been developed, but tend to rely on strong assumptions and unrepresentative samples. This paper extends the

extant literature by taking a counterfactual approach using propensity score matching, conducting a sensitivity analysis employing the Rosenbaum bounds to address selection bias on unobserved covariates, and using data from Add Health. Results show that while teen mothers' preexisting socioeconomic disadvantages and their lower level of cognitive and noncognitive skills play a nontrivial role, teen motherhood has significantly negative effects on early socioeconomic outcomes with the exception of public assistance receipt. The sensitivity analysis suggests that selection bias due to unobserved covariates would have to be very powerful to nullify these findings.

ACKNOWLEDGEMENTS

I have had ample opportunity to work with a number of extraordinary people at UNC. This dissertation would not have been written without their guidance and support. First of all, I am truly grateful to Kathleen Mullan Harris, my advisor. Kathie's unwavering faith in my capabilities, brilliance in seeing the big picture without overlooking the details, and sharp but constructive criticism of all of my scholarly work enabled me to grow as a sociologist. At every stage of my graduate career, she has been a constant support for me with her warmth, patience, and candor. I feel immeasurably fortunate to have had this splendid person as my mentor and friend.

I would also like to thank members of my dissertation committee. Guang Guo's insightful comments on my dissertation have deepened my understanding of theoretical and methodological issues with respect to studying social inequality. Philip Cohen has taught me how to think sociologically, and his advice has been indispensable for me to theorize and operationalize the role of skill differences in socioeconomic inequality. I would like to thank Ted Mouw for his robust support and encouragement. His rigor, versatility, and passion for social demography inspired me to make the best effort to improve my dissertation. I owe special thanks to Krista Perreira for giving me invaluable research opportunities. Not only has she guided my intellectual development through numerous conversations and comments on my research, but she has also provided moral support with her optimism whenever I needed a lift.

My fellow graduate students and friends have enriched my years in Carolina. I want to thank Hedy Lee, Jessica Hardie, Cheol-Sung Lee, Daniel Adkins, Tyson Brown, Seulki Choi, Hak-Seon Lee, Yong-Kyun Kim, and Phil Kim for their thoughtful suggestions and intellectual stimulation. I also wish to acknowledge the institutional support of the Carolina Population Center (CPC). I am especially grateful to Kathie Harris for serving as my preceptor, Barbara Entwisle for her valued advice and encouragement, and Jan Hendrickson-Smith for providing CPC pre-doctoral trainees with an excellent research environment.

The support of my family has been consequential to me since I decided to be a sociologist as my profession. Far in distance but close at my heart, my parents, Won-Il and Sun-Hwa, and my sister, Myung-Joo, in Korea have given me their endless love and prayer.

I would like to express my deepest gratitude to my wife, Swan. She has stood by me as my biggest supporter, my colleague, and my soul mate. I cannot thank her enough for all she has done for me, and truly wish her academic talent to reach its fullest potential. Finally, I am most blessed to have our adorable daughter, Haeon. For a spark and joy she brings to our life, I thank you.

TABLE OF CONTENTS

LIST OF TABLES x

LIST OF FIGURES xi

Chapter

I. INTRODUCTION 1

Previous Research 2

The Role of Cognitive and Noncognitive Skills in Socioeconomic Inequality 5

Specific Aims of the Dissertation 8

Overview of the Dissertation 8

II. THE ECONOMIC RETURN TO EDUCATION REVISITED: THE ROLE OF COGNITIVE AND NONCOGNITIVE SKILLS IN WAGE INEQUALITY 11

Introduction 11

Wage Inequality and the Education-Labor Market Nexus 13

The Role of Cognitive and Noncognitive Skills in Between- and Within-Education Group Wage Inequality 17

Research Questions 23

Data, Measures, and Method 25

Results 32

Discussion and Conclusion 40

III. EDUCATIONAL ASSORTATIVE MATING IN ADOLESCENCE AND ADULTHOOD: THE ROLE OF COGNITIVE AND NONCOGNITIVE SKILLS 51

Introduction 51

Previous Research on Educational Assortative Mating 54

Cognitive and Noncognitive Skills, Educational Assortative Mating, and Socioeconomic Inequality 58

Data, Measures, and Methods 62

Results 72

Discussion and Conclusion 81

IV. THE EARLY SOCIOECONOMIC EFFECTS OF TEENAGE CHILDBEARING: A PROPENSITY SCORE MATCHING APPROACH 94

Introduction 94

Socioeconomic Effects of Teenage Childbearing: Is It Causal? 96

A Counterfactual Approach: Propensity Score Matching 100

Data and Measures 106

Results 112

Discussion and Conclusion 120

V. CONCLUSION 131

Review of the Findings 131

Implications of the Research 135

Limitations of the Research and Future Directions 137

APPENDIX A: The Rosenbaum Bounds Method 140

APPENDIX B: Covariate Balance Check 142

REFERENCES 143

LIST OF TABLES

Table

Table 2.1. Descriptive Statistics by Educational Attainment, NLSY79 Sample 45

Table 2.2. Parameter Estimates from Quantile Regression Models for Log Hourly Wages, NLSY79 2002 Sample 46

Table 2.3. Percent Reduction in Wage Inequality within College Graduates by Skill Differences, NLSY79 2002 Sample 47

Table 2.4. Parameter Estimates from Quantile Regression Interaction Models for Log Hourly Wages, NLSY79 2002 Sample 48

Table 3.1. Means and Proportions of Analysis Variables by Educational Assortative Romantic Relationships among Adolescents, Add Health 84

Table 3.2. Multinomial Logistic Regression of Educational Assortative Romantic Relationships for Girls, Add Health 85

Table 3.3. Multinomial Logistic Regression of Educational Assortative Romantic Relationships for Boys, Add Health 87

Table 3.4. Means and Proportions of Analysis Variables by Education-Based Marital Sorting among Adults, NLSY79 89

Table 3.5. Multinomial Discrete-Time Event History Analysis of Education-Based Marital Sorting for Women, NLSY79 90

Table 3.6. Multinomial Discrete-Time Event History Analysis of Education-Based Marital Sorting for Men, NLSY79 92

Table 4.1. Descriptive Statistics by Teenage Childbearing Status: Full Sample 124

Table 4.2. Parameter Estimates from Logit Model Predicting Teenage Childbearing Status 125

Table 4.3. Propensity Score Matching Estimates of the Socioeconomic Effects of Teenage Childbearing 126

Table 4.4. A Sensitivity Analysis Using the Rosenbaum Bounds of the Causal Effects of Teenage Childbearing 127

Table 4.5. Propensity Score Matching Estimates of the Effects of Teenage Childbearing, by Race/Ethnicity 128

Table B1. Covariate Balance Check and Absolute Bias Reduction: Matched Sample 142

LIST OF FIGURES

Figure

Figure 2.1. Evolution of the College Wage Premium, NLSY79 1990-2002 Sample.. 49

Figure 2.2. Evolution of Wage Inequality within College Graduates, NLSY79 1990-2002 Sample 50

Figure 4.1. Predicted Probability of Teenage Childbearing: Full Sample.............. 129

Figure 4.2. Predicted Probability of Teenage Childbearing: Matched Sample 130

CHAPTER 1

INTRODUCTION

This dissertation examines the effect of cognitive and noncognitive skills on three socioeconomic outcomes: wage differentials, individual patterns of education-based mate selection, and the socioeconomic consequences of teen motherhood. Explaining socioeconomic inequality has been a backbone of sociological research with an emphasis on incorporating social structure along with individual characteristics in analytical models (Morris and Western 1999; Neckerman and Torche 2007). Traditionally, trend analyses and status attainment models have been predominant in the literature on socioeconomic inequality. Trend analyses clarify the ways in which basic demographic factors are associated with socioeconomic inequality over time, taking compositional changes in those factors into consideration. Status attainment models elaborate the mechanisms by which levels of socioeconomic inequality increase or decrease with a focus on intergenerational mobility. Certain micro-level determinants of differentials in socioeconomic attainment, however, have been overlooked in both types of approaches: personal traits and habits are regarded as too individual-specific to have theoretical and empirical relevance to socioeconomic inequality.

On the other hand, one stream of the research has a particular interest in early predictors of socioeconomic outcomes, such as those observed in childhood and adolescence that are closely associated with various background factors. And yet, linkages between early correlates and subsequent socioeconomic attainment have not been identified in a systematic manner. More often than not, adolescents' own characteristics tend to be 1) substituted with family, school, and neighborhood contexts, 2) considered a prior indicator of the same outcomes measured in adulthood (e.g., youth employment as a predeterminant of adult employment), or 3) included as outcome-specific attitudinal constructs (e.g., aspirations and expectations) in an ad hoc way.

In this project, I extend this line of research by identifying cognitive and noncognitive skills formed during childhood and adolescence as a micro basis of socioeconomic inequality and by demonstrating that individual-level skill differences are an important channel through which intergenerational mobility is associated with various forms of socioeconomic inequality. This dissertation illuminates the processes by which individuals' own attributes interact with social structural opportunities and constraints to impact socioeconomic outcomes and evaluates the significance of skill differences to the (re)production of socioeconomic inequality. I expect this study to contribute to deepening our understanding of the causes and effects of socioeconomic inequality in the United States.

PREVIOUS RESEARCH

Wage Inequality

Scholarship on labor market inequality has witnessed rising wage inequality since the 1980s (Autor, Katz, and Kearney 2005; Juhn, Murphy, and Pierce 1993; Morris and Western 1999). It has identified the economic return to education as a major source of changes in and levels of wage inequality. The college wage premium is considered a parsimonious but comprehensive account for the increase in overall wage inequality, because it captures the interaction between the demand for skills and the supply of skills through human capital investment. However, some critical concerns intrinsic to studying wage inequality require a refinement of the explanation based on the economic return to education. Those concerns include the importance of within-education group wage inequality (Bernhardt et al. 2001; Buchinsky 1998), the polarization of the current U.S. wage structure (Autor, Katz, and Kearney 2006; Lemieux 2006), and the unsettled issue of the causal effect of education on wage differentials (Angrist and Krueger 1991; Card 1995a; Herrnstein and Murray 1994). Some researchers have investigated the potential of unobserved skill bias, which is conventionally captured by cognitive skills, to illuminate the education-labor market nexus (Belzil and Hansen 2002; Murnane, Willett, and Levy 1995). Still, this line of research has not provided satisfactory explanations for why between-education group wage inequality accounts for only about one third of the total increase in overall wage inequality, why the upper-half wage inequality contributes greater to overall wage inequality, and what sources other than cognitive skills cause within-education group wage inequality. To better understand overall wage inequality, therefore, my dissertation examines whether cognitive and noncognitive skills explain between- and within-education group wage inequality at the same time.

Educational Assortative Mating

Extensive research to date on marriage has pointed out a growing association between socioeconomic inequality and increases in educational assortative mating in the United States (Fernández, Guner, and Knowles 2005; Kalmijn 1998; Mare 1991). The prevalence of the educational resemblance of spouses in union formation may indicate social closure, which reflects a further decrease in interactions between social status groups. Previous research has shown that as who marries whom depends increasingly on educational attainment and earnings potential, educational assortative mating has increased over the past six decades even after taking demographic change in the education distribution into account (Schwartz and Mare 2005). However, there has been little research on how differentials in social, cultural, and economic resources lead to educational assortative mating. Although education has been offered as a major explanation for these linkages, it is problematic because while background factors have an indirect impact on one's education-based partner choice by their influence on his or her educational attainment, education also forges preferences for cultural similarity and provides opportunities for mate selection by functioning as local marriage markets (Kalmijn 1998). In order to differentiate these confounding roles of education, we need to uncover unidentified sources that would clarify the relation between educational assortative mating and socioeconomic inequality. Moreover, most research has yet to incorporate the intimate nature of mating process into its empirical analysis. While it is well recognized that both attitude/behavior-based fellowship and education-based mate selection are a form of homophily (McPherson, Smith-Lovin, and Cook 2001), we know strikingly little about the role of personal traits and habits in individual patterns of educational assortative mating. This lack of attention in attitudinal and behavioral characteristics may be due to the assumption that they are too idiosyncratic and weakly correlated with status to predict education-based partner choice (Kalmijn 1994). However, given that some of these traits (e.g., intelligence and self-esteem), are developed unevenly according to family socioeconomic status, it is necessary to examine whether these personal traits and habits function as an important link between family background and mate selection by education.

Socioeconomic Effect of Teenage Childbearing

Despite the adverse life cycle consequences of teen motherhood (An, Haveman, and Wolfe 1993; Hofferth and Hayes 1987), it is still theoretically unclear whether the negative outcomes among teen mothers result from the incidence of childbearing *per se* or from the socioeconomic disadvantages they faced during childhood and adolescence (Hoffman 1998). While human capital theory holds that teenage childbearing has an exogenous effect on its socioeconomic outcomes given that it directly interferes with adolescent investment in human capital (Becker 1993), the revisionist view claims that teenage childbearing has an endogenous effect because it occurs mostly among disadvantaged female adolescents (Geronimus, Korenman, and Hillemeier 1994). Isolating the effect of teen motherhood creates a considerable methodological challenge known as omitted variables and selection biases (Winship and Mare 1992; Winship and Morgan 1999). If both observed and unobserved preexisting characteristics of teen mothers account for the relationship between teenage childbearing and its socioeconomic consequences, any assertion of causality becomes vulnerable. Because standard regression methods are not likely to be robust due to their failure to adequately control for preexisting socioeconomic differences between teen mothers and non-teen mothers, most of the alternative models have concentrated on finding better comparison groups (Cherlin 2001; Korenman, Kaestner, and Joyce 2001). For example, within-family fixed-effects models compare teen mothers with their sisters who gave birth after their teenage years to control for unobserved family-level heterogeneity. Quasi-natural experimental approaches take the approximate randomization procedures with observational data, treating twin births or miscarriages as comparison cases. Instrumental variables methods utilize variables that capture the exogenous component of teenage childbearing to mitigate the selection bias problem. Even if intuitively appealing, all of these models have their own drawbacks. It is not uncommon to find that they are grounded on somewhat strong assumptions and/or unrepresentative samples. Therefore, studies of the socioeconomic effect of teenage childbearing should take into account female teens' characteristics that are unmeasured in most of the prior research and apply a more flexible methodological approach that allows for identifying a reliable comparison group to teen mothers. My dissertation takes a counterfactual approach to address the causal effect of adolescent fertility, using propensity score matching.

THE ROLE OF COGNITIVE AND NONCOGNITIVE SKILLS IN SOCIOECONOMIC INEQUALITY

This dissertation explicitly seeks to fill the theoretical and analytical gaps in the literature on socioeconomic inequality described above by investigating the role of cognitive and noncognitive skills. Cognitive skills point to a general intelligence, or the "*g*" factor, whereas noncognitive skills represent attitudinal and behavioral personal traits and habits. These enduring dispositions include perseverance, self-confidence, sociability, emotional stability, interpersonal skills, and a future orientation, which are correlated to but distinct from cognitive skills (Farkas 2003). While both cognitive and noncognitive skills are highly dependent on changes in social conditions, noncognitive skills are more likely to be formed during childhood and stabilize during adolescence, compared to cognitive skills that are likely to be shaped during early childhood (Carneiro and Heckman 2003).

This study formulates noncognitive skills as a form of cultural capital in order to capture their important aspects with respect to socioeconomic outcomes. In elaborating Bourdieu's (1977) concept of *habitus*, Swidler (1986) defines cultural capital as a "tool kit" that constructs "strategies of action," such as accepted skills, styles, and informal know-how. Family background plays a crucial role in determining individuals' level of cultural capital, because they are not likely to be successful in socioeconomic outcomes without already having relevant cultural equipment (Lareau 2002). Moreover, cultural capital is easily convertible into economic and social capital and *vice versa* (Bourdieu 1984; Bourdieu and Wacquant 1992). Cultural capital can be utilized in accumulation of human capital in ways that are rewarded and reinforced in schools and workplaces (Bowles and Gintis 2002; Farkas et al. 1990; Rosenbaum 2001). Alongside cognitive skills, therefore, noncognitive skills as a form of cultural capital may have important consequences for socioeconomic outcomes. Consistent with these theoretical perspectives, numerous studies of cognitive and noncognitive skills have also provided empirical evidence for their positive impact on educational attainment and labor market outcomes (Jencks et al. 1979; Heckman, Stixrud, and Urzua 2006).

In this project, I explore how a focus on skill differences can improve upon the previous research on socioeconomic inequality. First, I argue that a close assessment of the role of cognitive and noncognitive skills illuminates the ways in which the principal-agent problem is dealt with in the labor market (Bowles, Gintis, and Osborne 2001). Given imperfect and asymmetric information inherent in employment relations, it is difficult for employers to detect employees' effort level with their known characteristics such as educational attainment. To resolve this problem, employers establish an incentive structure within the workplace. Employees who have higher levels of cognitive and noncognitive skills are more likely to respond positively to this incentive structure, and as a result, employers reward a wage premium to them. Thus, skill differences could contribute to both between- and within-education group wage inequality.

Second, I expect that a careful examination of cognitive and noncognitive skills will depict the ways in which familial resources affect education-based mate selection. Individuals' skill levels, once constructed as a pre-schooling, pre-labor market factor, may function as an important link between family socioeconomic status and patterns of educational assortative mating. In addition, skill differences may involve the interactive dynamics of mate selection. In this matching process, individuals' attitudinal and behavioral characteristics should have a direct relevance to forming long-term intimate relationships. Taken together, the association between cognitive and noncognitive skills and mate sorting by education may clarify how socioeconomic inequality plays out at an intimate level of mating process.

Third, I propose inclusion of cognitive and noncognitive skills to reduce the omitted variable problem that tends to make estimation of the causal effect of teenage childbearing biased. It is surprising that most studies do not take skill differences into account, because adolescents with higher levels of skills not only have a favorable family background but also are more likely to succeed in schools and labor markets. While controlling for cognitive and noncognitive skills would not clear up selection bias, it certainly helps to isolate the exogenous effect of early motherhood on subsequent socioeconomic outcomes.

In summary, by gauging the extent to which skill differences play a role in wage inequality, educational assortative mating, and the incidence of teenage childbearing, this dissertation will increase our knowledge about the mechanisms by which individual characteristics shaped in childhood and adolescence produce, and/or reproduce, socioeconomic inequality in light of intergenerational mobility.

SPECIFIC AIMS OF THE DISSERTATION

To investigate the role of cognitive and noncognitive skills in each domain of socioeconomic inequality, this project addresses the following three specific aims:

1) To examine a) the effect of skill differences on between- and within-education group wage inequality; b) the contingency of the wage effect of cognitive and noncognitive skills on the college premium; and c) changes in the role of skill differences in the college wage premium due to the aging and accumulation of labor market experience of college-educated workers.

2) To assess a) the association between cognitive and noncognitive skills and academic achievement-based romantic partner choice in adolescence; b) the effect of skill differences on education-based spouse selection in adulthood; and c) gender differences in the role of skill differences in individual patterns of educational assortative mating.

3) To apply a counterfactual approach to reexamine the causal effect of teenage childbearing on subsequence socioeconomic outcomes (educational attainment, labor market, and welfare outcomes) with introduction of cognitive and noncognitive skills.

OVERVIEW OF THE DISSERTATION

This dissertation is organized as three separate research articles. While all of these articles hold a common theme of the role of cognitive and noncognitive skills in socioeconomic inequality, each article examines their role in specific domains of socioeconomic outcomes:

wage differentials, mate sorting by education, and the socioeconomic consequences of teen motherhood.

Chapter 2 proposes an analysis of the impact of cognitive and noncognitive skills on between- and within-education group wage inequality, using data from the National Longitudinal Survey of Youth 1979 (NLSY79). Drawing on the literature that postulates the importance of skill differences in the labor market where incomplete information is chronic, I discuss why and how cognitive and noncognitive skills may play a role in explaining wage inequality. With an application of a quantile regression method (Buchinsky 1998; Hao and Naiman 2007; Koenker and Bassett 1978), this chapter simultaneously addresses how much influence cognitive and noncognitive skills have on between- and within-education group wage inequality in light of the increasing return to education.

Chapter 3 of my dissertation explores the extent to which individual patterns of educational assortative mating vary by cognitive and noncognitive skills, using data from the National Longitudinal Study of Adolescent Health (Add Health) and the NLSY79. Given that individuals' skill levels are highly dependent on familial resources, the close association between skill differences and mate sorting by education indicates a more sophisticated way in which intergenerational transmission of social, cultural, and economic resources reduces interactions between different status groups, resulting in an increasing social divide. Employing multiple regression and multinomial discrete-time event history analyses, I investigate whether individuals with high cognitive skills and socially valued noncognitive skills are more likely to date the academically achieved during adolescence and to marry the more educated during adulthood and how this association is moderated by gender.

In Chapter 4, I reevaluate the socioeconomic consequences of teenage childbearing. Using data from Add Health, this study extends the literature by examining the role of cognitive and noncognitive skills in the incidence of teenage childbearing and by employing a counterfactual approach using propensity score matching (Rosenbaum and Rubin 1983;

Rubin 1977). While unmeasured in most research on teenage childbearing, cognitive and noncognitive skills are found to affect early motherhood. The counterfactual analysis attempts to obtain the causal effects of teenage childbearing by identifying a better comparison group to teen mothers, based on observed covariates. A sensitivity analysis is conducted to address selection bias on unobserved covariates (Rosenbaum 2002). Add Health helps to capture recent changes in welfare policy and adolescent population composition. This chapter expects to give new insights into the relationship between teenage childbearing and its socioeconomic consequences.

CHAPTER 2

The Economic Return to Education Revisited:
The Role of Cognitive and Noncognitive Skills in Wage Inequality

INTRODUCTION

Rising wage inequality since the 1980s has drawn much attention in the social sciences (Autor, Katz, and Kearney 2005; Juhn, Murphy, and Pierce 1993; Morris and Western 1999). Social scientific research has proposed a number of factors that can explain well-known changes in and levels of wage inequality. There have been increasing demand for skills due to macro-level economic changes, such as industrial restructuring, institutional reconfiguration, and technological changes, and shifts in the labor supply corresponding to the labor demand, such as compositional changes in educational attainment and the strengthened correlation between workers' skill levels and their highest education completed. Economic literature on wage inequality argues that the economic return to education captures the process by which the demand for skills and the supply side response to it produce a growth in wage inequality and its high level observed in the current U.S. labor market (Katz and Murphy 1992; Murphy and Welch 1992). The college wage premium epitomizes the tightened nexus between education and labor market outcomes.

Research focusing on the economic return to education, however, tends to overlook at least three aspects of changes in and levels of overall wage inequality. First, it is basically an account based on between-education group wage inequality. Although wage differentials between education groups are important in their own right to understand the sources of overall wage inequality, less attention has been paid to within-education group wage inequality, which is defined as wage dispersions among workers with similar education and labor market experience (Bernhardt et al. 2001; Buchinsky 1998; McCall 2000). Second, the recent literature on wage inequality has found the so-called polarization of the U.S. wage structure (Autor, Katz, and Kearney 2006; Lemieux 2006). Wage inequality has relatively stabilized in the lower portion of the wage distribution over the last 30 years, whereas it has continued to grow in the upper portion of the wage distribution. Studies of wage inequality need provide potential explanations for this principal aspect of the current U.S. wage structure. Third, much research questions the causal role of education in wage determination (Belzil and Hansen 2002; Herrnstein and Murray 1994; Juhn, Murphy, and Pearce 1993; Murnane, Willett, and Levy 1995; Taber 2001). It posits that the demand for cognitive skills is an underlying factor of the economic return to education. Because individuals with higher cognitive skills are more likely to choose higher education, differences in workers' educational attainment may simply be a reflection of differences in unobserved ability among workers. However, other research has shown that even taking cognitive skills into account does not fully explain the effect of education on wages (Angrist and Krueger 1991; Ashenfelter and Rouse 1998; Card 1995a). These contrasting findings suggest that there still remain other unidentified sources of between- and within-education group wage inequality.

To fill these gaps, this paper seeks to provide a more elaborated explanation for the economic return to education by examining the role of both cognitive and noncognitive skills in between- and within-education group wage inequality. Despite the importance of within-education group wage inequality to overall wage inequality, the polarized pattern of the current U.S. wage structure, and the unsettled issue of the causal relationship between education and wages, sociological research on wage inequality to date has been relatively scant (Morris and Western 1999; Neckerman and Torche 2007).[1] On a theoretical front, this paper emphasizes the multidimensionality of skills. Drawing on the literature on forms of

[1] Kim and Sakamoto (2008) and Mouw and Kalleberg (2008) are two of the few sociological studies that address both between- and within-group wage inequality but their analytical focus is on occupational structure.

capital, I conceptualize noncognitive skills as a form of cultural capital that is referred to as attitudinal and behavioral personal traits and habits (Bourdieu 1984; Lareau 2002; Swidler 1986). Research shows that correlated with but distinct from cognitive skills, noncognitive skills are highly contingent on differences in social, economic, and cultural resources (Carneiro and Heckman 2003; Farkas 2003). In addition, noncognitive skills are considered a critical factor that resolves employer-employee mismatches that are prevalent in the labor market (Bowles, Gintis, and Osborne 2001). A growing body of literature has documented that both cognitive and noncognitive skills are valued in labor markets (Bowles and Gintis 2002; Heckman, Stixrud, and Urzua 2006).

On a methodological front, I use data from the National Longitudinal Survey of Youth 1979 (NLSY79) and employ a quantile regression method to overcome the drawbacks of ordinary least squares (OLS), a dominant method in the literature on wage inequality (Hao and Naiman 2007; Koenker and Bassett 1978). Because OLS estimates the conditional mean effects of the determinants of wages, it is useful to detect the sources of between-group wage inequality but makes it difficult to evaluate the role wage determinants play in within-group wage inequality. By assessing the relative contribution of cognitive and noncognitive skill differences to the college wage premium at various locations of the wage distribution (e.g., at the lower and upper tails and the median), the quantile regression method simultaneously addresses between- and within-education group wage inequality.

WAGE INEQUALITY AND THE EDUCATION-LABOR MARKET NEXUS

One of the well-known recent economic trends in the United States is a rise in wage inequality. Overall wage inequality rapidly increased in the 1980s and though relatively stabilized in the 1990s, remains at a historically high level (Autor, Katz, and Kearney 2005). Social scientific research has provided various explanations for changes in and levels of wage inequality (see Neckerman and Torche 2007 for a summary). First, there are labor demand-side factors that include industrial restructuring, institutional reconfiguration, globalization, immigration, and technological changes. For example, deindustrialization, globalization, and immigration reduce demand for mid-level jobs concentrated in manufacturing sector, while increase demand for low-level jobs prevalent in service sector. Institutional reorganization results in demand for high-ranking managerial jobs such as CEOs. Technological changes like computerization in the workplace give rise to demand for high-skilled jobs. Second, labor supply-side factors include the composition of workers by educational attainment, school quality, and changes in and levels of human capital investment. For instance, if a worker's educational attainment becomes more diverse and there is increasing demand for the educated, wage inequality is likely to increase. The economic return to education has been found to be a critical factor that captures both the demand for skills and the labor supply-side responses to it.[2] As Card (1995b:23) put it, "one of the most important 'facts' about the labor market is that individuals with more education earn higher wages." The college wage premium is evidence for the tightened education-labor market nexus.

However, most accounts of the economic return to education are concentrated on between-education group wage inequality, which is summarized by wage differentials in means between education groups. Research that attempts to explain overall wage inequality reveals that the increasing economic return to education accounts for only about one third of the total increase in wage inequality (Bernhardt et al. 2001), indicating that within-education group wage inequality occupies a larger portion of overall wage inequality (Autor, Katz, and

[2] This paper does not take into account all these labor demand and supply side factors to address the sources of wage inequality. The focus here is on potential sources of the economic return to education at the individual level—workers' characteristics and responses to shifts in the labor demand for skills and changes in the labor supply. Most of the macro-level factors, therefore, are treated as given in this study. For a detailed discussion on these factors, see Morris and Western (1999); Neckerman and Torche (2007).

Kearney 2005).[3] Indeed, wage dispersions have become largest among college graduates, with the smallest wage dispersions observed among high school dropouts (Buchinsky 1998; McCall 2000). Moreover, a close look at the trend in wage inequality shows that the upper-half overall wage inequality steadily rose while the lower-half overall wage inequality ceased to rise since the late 1980s (Autor, Katz, and Kearney 2006; Lemieux 2006). This finding identifies a polarization of wage growth, where the most rapid rise occurred in the top quartile of the wage distribution and slower wage growth in the middle two quartiles than in the bottom quartile. The polarized pattern of wage growth during the 1990s implies that the economic return to education may not be uniform across the wage distribution. Therefore, it is important to specify what factors contribute to the differential effects of educational attainment at various locations of the wage distribution. To address this concern, research needs to examine both between- and within-education wage inequality in order to better understand patterns of overall wage inequality.

Yet discussions based on between-education group wage inequality tend to regard within-education group wage inequality as residuals, assuming that this is attributable to differences in unobserved skills, especially cognitive skills.[4] There has been much reluctance to accept the causal effect of education on wage inequality, because education may not be truly exogenous to wage differentials given that it involves individuals' schooling choice. As macroeconomic changes have altered the wage structure since the late 1970s, the labor

[3] One possible explanation for within-education group wage inequality concerns the composition effect, which is an increase of college graduates in the labor market. However, Autor, Katz, and Kearney (2005) show that within-education group wage inequality has remained prevalent during the 1990s after taking the composition effect into account. Rather, the composition effect played a role in the stagnation of lower tail wage inequality since the late 1980s, offsetting the impact of the economic return to education in the lower tail of the wage distribution.

[4] Studies focusing on unobserved ability bias do not confine their interest to cognitive aspect of skills. Implicitly though, they recognize the wage effect of noncognitive skill. And yet, there has been little research that empirically assesses the relation of both skills to between- and within-education group wage inequality at the same time.

demand for skills has increased.[5] Because individuals who possess a higher level of cognitive skills are more likely to choose higher education, the effect of education on wages may simply represent unmeasured abilities of the more educated. In this scenario, what one's educational attainment signals to the labor market is his or her level of productivity-enhancing skills that are formed before schooling rather than education itself. Some research interprets growing residual wage inequality even after accounting for education and work experience as a sign of increases in the demand for skills (Juhn, Murphy, and Pearce 1993). A large body of economic literature has provided findings in favor of the unobserved ability bias story: When cognitive ability measured by standardized test scores is taken into account or unobserved ability is allowed to correlate with both educational attainment and wages, the effect of education on wages either reduces to a nontrivial extent or becomes insignificant (Belzil and Hansen 2002; Murnane, Willett, and Levy 1995; Taber 2001). Herrnstein and Murray's (1994) controversial study goes further to rule out the role of education in labor market outcomes, maintaining that differences in intelligence, which they treat as genetically inherited, account for most of the racial and gender gaps in wages.

However, how much of the economic return to education is due to unobserved ability bias is not settled (Cawley et al. 2000). Unlike the research supporting the unobserved ability bias story, refined human capital models have provided evidence for the causal effect of education on wage differentials (See Card (1995b, 1999) for a review). Numerous economic and sociological studies found that the effect of schooling is significant after controlling for cognitive ability, family background, and school quality. This finding holds true even as more rigorous analytical approaches are used to control for unobserved heterogeneity. Whether employing instrumental variables method (Card 1995a; Angrist and Krueger 1991) or exploiting sibling or twin data (Ashenfelter and Rouse 1998), both approaches produce a

[5] These macroeconomic changes include expansion of the service sector along with decline in the manufacturing sector, erosion of minimum wage, deunionization, and computerization leading to skill-biased technological changes (SBTC).

larger effect of education on wage differentials than does a conventional OLS method. These findings imply that while between-education group wage inequality can result in part from differences in cognitive ability, other sources of between-education group wage inequality should be also identified. Furthermore, most interest in this line of research lies in how much of the economic return to education, not how much of the wage dispersions within education group, is explained by taking cognitive skills into account.

In the following section, this paper investigates the effect of cognitive and noncognitive skills on both between- and within-education group wage inequality. Apart from cognitive skills, noncognitive skills may function as an unidentified source of wage differentials between education groups because noncognitive skills are likely to affect individuals' education decision. Furthermore, noncognitive skills may play a role in large wage dispersions among individuals with similar education, given that employers' demand for skills is not simply confined to cognitive skills but requires employees to internalize the norms of attitudes and behaviors set in the organization of workplace.

THE ROLE OF COGNITIVE AND NONCOGNITIVE SKILLS IN BETWEEN- AND WITHIN-EDUCATION GROUP WAGE INEQUALITY[6]

Theoretical Perspectives

Noncognitive skills are referred to as attitudinal and behavioral personal traits and habits. Whereas cognitive skills point to a general intelligence, or the "*g*" factor, usually measured by standardized test scores and occasionally course grades, noncognitive skills pertain to enduring dispositions that are not captured by cognitive skills and highly dependent on changes in social conditions (Carneiro and Heckman 2003). Unlike cognitive skills, noncognitive skills are attributable to an infinite number of traits and habits, so research has

[6] Since cognitive skills have been disproportionately utilized as a reliable proxy for individual skill differences, this section focuses more on the relation of noncogntive skills to labor market outcomes. For a review of the role of cognitive skills in wage inequality, see Farkas (2003).

assessed noncognitive skills in light of its relevance to socioeconomic success, such as perseverance, self-confidence, motivation, sociability, emotional stability, interpersonal skills, and a future orientation (Farkas 2003).[7] From a developmental perspective, noncognitive skills are more likely to be shaped during childhood and stabilized during adolescence, compared to cognitive skills that are likely to be formed during early childhood (Bowles and Gintis 2000; Heckman and Rubinstein 2001).[8]

The idea of noncognitive skills can be better understood as a form of cultural capital. Its important aspect is addressed in Swidler (1986)'s extension of Bourdieu's (1977) concept of *habitus*. She considers culture as a "tool kit" that constructs "strategies of action": "One can hardly pursue success in a world where the accepted skills, styles, and informal know-how are unfamiliar. One does better to look for a line of action for which one already has the cultural equipment" (Swidler 1986:275).[9] In this vein, the conceptualization of noncognitive skills as a form of cultural capital reflects the importance of family background. Studies of cognitive and noncognitive skills show that these skills are unevenly distributed across children whose families differ by levels of socioeconomic resources and parenting practice (Condron 2007; Lareau 2002). In an ethnographic study of middle and working class families, Lareau (2002) observes that middle class children take a cumulative advantage of "concerted cultivation" in developing their cognitive and noncognitive skills. They gain a sense of

---

[7] In the literature, noncognitive skills are used interchangeably with soft skills or socioemotional skills, and measured by locus of control, self-esteem, externalizing/internalizing problem behaviors, and executive functioning among others. The term "noncognitive skills" has been adopted to distinguish soft skills from hard skills, which are measured by standardized test scores, although there is no doubt that soft skills also involve cognitive processes (Farkas 2003; Kuhn and Weinberger 2005). Thus, the term "noncognitive skills" is used throughout this paper mainly because this kind of skills has not been closely addressed in most research on the economic return to education, which primarily regards unobserved ability bias as a result of cognitive skills.

[8] According to Carneiro and Heckman (2003), cognitive skills are fairly stable after age 8, whereas noncognitive skills can be improved until the late teenage years.

[9] DiMaggio (1982) defines cultural capital more formally as interest in and experience with prestigious cultural resources. While highlighting an important dimension of cultural capital, this definition is not as directly concerned with personal traits and habits as is Swidler's.

entitlement through wide-ranging family resources, organized leisure activities, and extensive reasoning by their parents. Meanwhile, working class and poor children display a sense of constraint and powerlessness as a result of lack of family resources, their parents' belief in an "accomplishment of natural growth," and their frequent use of directives.

Another aspect of cultural capital is its convertibility into economic and social capital and *vice versa* (Bourdieu 1984; Bourdieu and Wacquant 1992). It is posited that as education takes a leading role in social stratification system, cultural capital can be utilized in accumulation of human capital in ways that are rewarded and reinforced in schools and workplaces (Bowles and Gintis 2002; Farkas et al. 1990; Rosenbaum 2001). Taken together, these aspects of noncognitive skills as a form of cultural capital have important implications for explaining wage inequality in light of intergenerational mobility, because they highlight that socialization for work takes place well before entering the labor market.

Why does noncognitive skills, which are not usually thought of as "skills," matter to wage inequality? Bowles, Gintis, and Osborne (2001) provide a theoretical framework with three alternative labor market models in which individuals' wage level is determined: 1) a "Walrasian" model; 2) a "Schumpeterian" model; and 3) a "Coasean" model.[10] In a Walrasian world, labor markets are assumed to rapidly reach equilibrium, which means that there is sufficient information that circulates on both the labor demand and supply sides. Employees' labor efforts are rewarded according to their level of productivity that is determined by their pre-market attributes. Hence, wage differentials entirely result from cognitive skill differences, i.e., productivity-enhancing skills, between individual workers in this neoclassical labor market model. However, in a Schumpeterian world in which the assumption of labor market equilibrium is not realistic, one should consider "disequilibrium

[10] Bowles, Gintis, and Osborne (2001) coin these terms because each model represents these economists' specification of earnings determination: Leon Walras offered a neoclassical model of labor market; Joseph Schumpeter introduced the concept of disequilibrium rents; and Ronald Coase raised the "principal-agent" problem that is prevalent in labor markets.

rents” as a determinant of wage differentials along with cognitive skill differences. These rents are generated by technological and organizational changes and other market shocks. Given that individuals’ response to the disequilibrium rents differ in their ability, those who are capable of capturing these rents are more likely to succeed in the labor market. Noncognitive skills such as self-directedness, internality as opposed to fatalism, and a future orientation are critical to dealing with disequilibria, although those personal traits might not be considered productivity-enhancing skills.

In a Coasean world, employees’ noncognitive skills are treated as a substantial element in wage determination processes. Employers are constantly faced with the “principal-agent” problem that stems from incomplete and asymmetric information: they cannot directly discern employees’ effort level as assumed in the Walrasian model, even if knowing employees’ educational attainment and cognitive ability. To motivate employees, thus, employers set the incentive structure within the workplace. For example, if monitoring employees’ work activities is not feasible, employers are more likely to value employees’ self-monitoring that leads to productivity in an indirect manner. This sort of personal traits is referred to as “incentive-enhancing preferences,” which include an orientation toward the future, personal efficacy, and internalized locus of control. Bowles, Gintis, and Osborne (2001) argue that given that cognitive skills do not fully account for the economic return to education, the Coasean (and Schumpeterian) model is a more plausible depiction of labor markets where employers reward a wage premium to employees with such noncognitive skills who positively respond to the employer-setting incentive structure. In this sense, noncognitive skills may function as an important determinant of not only the increasing economic return to education (i.e., between-education group wage inequality) but also large wage dispersions among individuals with similar education (i.e., within-education group wage inequality).

Empirical Findings

In support for these theoretical perspectives, recent literature has documented empirical findings of the role of noncognitive skills in labor market outcomes. One line of research directly examines what employers want from job applicants and employees; and another line focuses on the direct impact of cognitive and noncognitive skills in wage determination. First, unlike most of the research that is based on the labor supply side, Holzer's (1996) study utilizes an employer survey supplement to the Multi-City Study of Urban Inequality in order to address employers' hiring decisions for low-educated workers. He finds that both cognitive and noncognitive skills are a substantial part of what employers require of workers for task performance, as shifts in industries and occupations are more oriented toward "information-processing" jobs. Even in non-college and non-white-collar jobs, employers seek workers with high levels of cognitive and social skills. According to Holzer's estimates, employers interviewed nearly 90 percent of job applicants to make their personal hiring judgment in addition to available objective measures, with politeness and motivation the most highly stressed factors for hiring, followed by verbal skill and physical appearance.[11]

More importantly, employers' demand for noncognitive skills is not merely limited to the low-educated. A recent report of employers who hired new college graduates indicates that communication, motivation, teamwork, and leadership skills all have more influence on hiring decisions than do academic achievement or grade point average (National Association of Colleges and Employers 2000). In addition, in the Employers' Manpower and Skills Practices Survey, which is more general but conducted in Britain, personnel managers presented poor attitude, motivation, and personality as a major recruitment problem, compared to lack of technical skills (Green, Machin, and Wilkenson 1998). Taken together,

[11] Another employer survey also reveals that for employers seeking non-supervisory or production workers, the rank on the importance of workers' characteristics is attitude, communication skills, industry-based skill credentials, years of schooling, and academic performance in order (Bureau of the Census 1998).

employer surveys collected in the 1990s suggest that noncognitive skills are as much valued as cognitive skills in the employers' hiring decision processes.

Second, a significant body of literature on wage differentials has provided evidence for the wage effect of cognitive and noncognitive skills. Jencks et al. (1979) find that noncognitive skills such as perseverance, industriousness, and leadership had a positive effect on wages even after controlling for a number of human capital variables and cognitive skills, with noncognitive skills showing the larger effect than cognitive skills. In an analysis of the General Educational Development (GED) certificate, Heckman and Rubinstein (2001) demonstrate that GED recipients have higher cognitive skills than high school dropouts but earn lower wages than high school graduates, because they have lowest noncognitive skills among education groups. Their study suggests that the GED is a "mixed" signal in which the economic return to education is not simply reduced to cognitive skills and educational credentials.

Dunifon and Duncan (1998) show that motivational traits are as powerful a predictor as completed schooling of earnings differentials and of participating in on-the-job-training. Interestingly, their results also show that an orientation toward challenge and a sense of personal control have a stronger effect on earnings at later ages, suggesting that it takes a substantial amount of time for employers to assess workers' noncognitive skills. Dunifon, Duncan, and Brooks-Gunn (2001), for instance, show that "home cleanliness" is a significant factor predicting earnings outcomes 25 years later when controlling for socioeconomic background, own education and cognitive ability, and time spent in housework. They conjecture that home cleanliness reflects an overall ability to maintain a sense of order, which in turn carries over the ability to keep a degree of organization and efficiency that may be a skill valued in the labor market. Rosembaum (2001) finds leadership measured in high school strongly predicts earnings 10 years later, coupled with cognitive skills and other types of noncognitive skills. Kuhn and Weinberger (2005) report the relatively short-term, but significant, effect of leadership skills on wages. They also suggest that the leadership effect estimated in the 1990s is much larger than that in the 1970s.[12] Heckman, Stixrud, and Urzua (2006) confirm the studies described above, demonstrating that noncognitive skills constructed as a latent factor is as important as cognitive skills in explaining differentials in wages, employment status, and occupational attainment.

While all these studies clearly lead to the better understanding of between-education group wage inequality (i.e., why wages differ according to educational attainment), we still do not know much about the role of cognitive and noncognitive skills in within-education group wage inequality (i.e., why wage dispersions are considerable in groups with similar education). This paper extends the current literature by applying a quantile regression method to examine the role of cognitive and noncognitive skills in wage differentials both between education groups (high school graduates vs. college graduates) and within college graduates.

RESEARCH QUESTIONS

Previous research has proposed the economic return to education as a major source of wage inequality. Despite the unquestionable importance of between-education group wage inequality, recent changes in and levels of wage inequality also point to the role of within-education group wage inequality and the polarization of the U.S. wage structure in overall wage inequality. Most research on unobserved ability bias has focused heavily on cognitive skill with little interest in these current patterns of wage inequality. Drawing on the recent research on the role of cognitive and noncognitive skills in wage determination, this analysis addresses three research questions:

Question 1: Does taking cognitive and noncognitive skills into account reduce the effect of the college premium on wages and wage inequality within college graduates?

---

[12] Kuhn and Weinberger's (2005) study is an exception to most research on noncognitive skills and labor market outcomes in that they stress that leadership skills operate within managerial occupation, which implies the role of noncognitive skills in within-group wage inequality.

This question concerns the role of cognitive and noncognitive skills in between- and within-education group wage inequality. In the analysis, I evaluate the extent to which the effect of the college premium across the wage distribution and wage dispersions among college graduates reduce after controlling for cognitive and noncognitive skills. To investigate whether a polarized pattern of the wage structure is due to cognitive and/or noncognitive skills, this analysis assesses where in the wage distribution the role of cognitive and noncognitive skills are more pronounced.

Question 2: How contingent is the wage effect of cognitive and noncognitive skills on the college wage premium?

To take a closer examination of the relative contribution of cognitive and noncognitive skills in accounting for between-education group wage inequality and wage inequality within college graduates, this analysis tests whether the effect of cognitive and noncognitive skills on wage differentials is independent of the college premium across the wage distribution or whether it operates mainly within college graduates. I address this question by gauging two interaction effects on wages, one between the college premium and cognitive skills and the other between the college premium and noncognitive skills.

Question 3: Does the role of cognitive and noncognitive skills in the college wage premium change as workers reach their prime ages in the labor market?

This question concerns the shapes of the effect of the college premium across the wage distribution according to the aging and accumulation of labor market experience of college-educated workers. Previous studies of the education-labor market nexus have shown that educational attainment has an impact on wage differentials regardless of workers' age and labor market experience (Card 1995b, 1999), suggesting that between-education group wage inequality is always present. However, levels of wage inequality among college graduates would be relatively low among young workers because college-educated workers are likely to have less labor market experience. If so, there would not be much room for the contribution of cognitive and noncognitive skills in explaining wage inequality within young college graduates. Meanwhile, wage inequality within college graduates is likely to increase among more experienced workers as college-educated workers age and accumulate their labor market experience. In that case, the role of cognitive and noncognitive skills in wage inequality within college graduates would become more salient as they reach their prime ages in the labor market.

DATA, MEASURES, AND METHOD

Data

This paper uses data from the National Longitudinal Survey of Youth 1979 (NLSY79). The NLSY79 is a nationally representative sample of 12,686 young men and women who were 14 to 22 years old when they were first surveyed in 1979 (Center for Human Resource Research 2002). These individuals were interviewed annually through 1994 and since then, have been interviewed on a biannual basis. Despite its long term coverage, retention rates are high. Since their first interview, many of the respondents have made transitions from school to work. These data make it possible to study a large sample that represents American men and women born in the late 1950s and 1960s, who have witnessed rising wage inequality over the last 30 years throughout their labor market experiences. A key feature of this survey is that it gathers information in a work history format, in which dates are collected for the beginning and ending of work-related experiences. Educational attainment and labor market performance are detailed in this manner. The NLSY79 provides rich sets of variables for earnings, respondents' cognitive and noncognitive skills, educational attainment and work experience, family background, and other key demographic characteristics.

The analytic sample is restricted to individuals who worked at least one week in the year prior to the survey year; who did not attend school when information on their wages was gathered; and who were among noninstitutional civilian population. The sample is further restricted to individuals who were 15-18 years old in 1980, because measures of cognitive and noncognitive skills are constructed as pre-schooling and pre-labor market factors to minimize reverse causality (see the measures section for more detailed description).[13] For Research Questions 1 and 2, the analysis is conducted with the 2002 sample because not only is it the most recent one available from the NLSY79 data but also respondents in this sample were at their prime working age. After constructing all measures, the final analytic sample size is 2,527. For Research Question 3, the analysis utilizes multiple biannual samples. Because college education delays labor market entry, it is likely that less-educated workers are overrepresented in earlier waves of the NLSY79. Therefore, I use the samples from 1990 to 2002. These multiple samples allow me to address whether and how the wage effect of cognitive and noncognitive skills varies by workers' aging and time spent in the labor market (Duncan and Dunifon 1998). While respondents were in the late 20s with the average labor market experience of about 8 years in the 1990 sample, they were in the late 30s and early 40s with the average labor market experience of about 18 years in the 2002 sample.

Measures

HOURLY WAGES The dependent variable of this study is log hourly wages averaged over 3 or 5 years.[14] It has been shown that wages are more reliable if averaged over multiple years than if measured only at one year. Wage data include respondents who had wage information for at least two of the 3 or 5 year span. Wage rates in 2000 dollars below $1 were set equal to

[13] This study cautions against the generalizability of results reported here, given limitations of the analytic samples such as age range and an underrepresentation of the white population (see Table 2.1). Current Population Survey (CPS) has been widely used in the wage inequality literature, but it does not contain information on cognitive and noncognitive skills. A preliminary analysis (not shown) suggests that there are some discrepancies in demographic compositions between the CPS and the NLSY79, but the basic patterns of the economic return to education are very similar in both data sets. Also, there is little substantial difference between the analytic samples used here and the samples before listwise deletion.

[14] For the samples from 1994 on, only three observations per individual are available because the NLSY79 has collected data biannually since then.

$1 and wage rates above $100 were set to $100. Nominal data are inflated to 2000 price levels by the implicit deflator of personal consumption expenditures for gross national product.

COGNITIVE AND NONCOGNITIVE SKILLS The variables of most interest in this study are levels of cognitive and noncognitive skills individuals possess during adolescence. For a measure of cognitive skills, the NLSY79 includes an aptitude indicator, the full Armed Services Vocational Aptitude Battery (ASVAB) consisting of a series of tests measuring knowledge and skill in areas such as mathematics and language. It was administered to 94 percent of the sample respondents in 1980. A composite score derived from the ASVAB is used to construct an Armed Forces Qualifications Test (AFQT) score, which has been extensively used to measure cognitive skills (Cawley et al. 2000).

Noncognitive skills are measured by several indexes consisting of attitudinal and behavioral aspects of such skills. The NLSY79 provides the Rotter's locus of control scale and the Rosenberg's self-esteem scale, administered in 1979 and 1980, respectively. Locus of control measures the degree of control individuals feel ranging from external to internal. According to Rotter (1966), individuals who believe that outcomes are due to luck have an external locus of control while individuals who believe that outcomes are due to their own efforts have an internal locus of control. The self-esteem scale measures perceptions of self worth (Rosenberg 1965). I use a composite index of both scales, which has been commonly employed in past research on the effects of noncognitive skills on socioeconomic outcomes (Carneiro and Heckman 2003).

This study constructs pre-schooling and pre-labor market measures of cognitive and noncognitive skills to reduce endogeneity by which unexpected academic and labor market success might develop both skills valued in schools and labor markets. Specifically, these measures are constructed by 1) restricting the analytic sample to respondents who were adolescents in 1980 (15 to 18 years old) and 2) calculating residual values from a regression model where the composite index of cognitive or noncognitive skills is regressed on age dummies and years of schooling in 1980. Then these residual values are standardized to serve as the measures of both skills. We need to be cautious in using these measures because there may be an issue of measurement error. It is unclear how perfectly both measures represent individuals' cognitive and noncognitive skills, but these measures resonate well with the theoretical views described earlier, and, as already noted, they have been used frequently in studies of cognitive and noncognitive skills. In addition to constructing these skills as the pre-schooling and pre-labor market factors, I introduce both skill measures simultaneously in the full model in order to cancel out some of the measurement error problem, given that motivated individuals are more likely to obtain higher test scores.

EDUCATIONAL ATTAINMENT Educational attainment is a key variable in this analysis. It is measured by highest level of education completed at the time hourly wages were measured. This consists of a set of mutually exclusive and exhaustive dummy variables, which are high school dropouts, high school graduates (reference group), some college attendees, and college graduates. This study treats GED holders as high school dropouts, following Cameron and Heckman (1993) and Heckman and Rubinstein (2001) that show GED holders are more similar to high school dropouts than high school graduates in various adult outcomes. In the analysis, I focus on the coefficients of being college graduates to address the relationship between the college wage premium and cognitive and noncognitive skills.

FAMILY BACKGROUND Family background covers family structure and parental education at age 14, and parental occupational status, the number of siblings, and family income in 1979. Family structure is categorized as two biological-parent families (reference group), two-parent step families, single-mother families, and other families (e.g., single father families or foster families). Parental education is measured with the highest level of education either of the parents obtained and has the same classification scheme as respondents' level of education. Parental occupational status is measured by whether or not either of the parents was in professional or management occupations, based on the 3-digit occupational classification.

CONTROL VARIABLES In addition to family background, this study includes other demographic variables such as sex (male/female), race/ethnicity, residence in the South and urban residence at age 14. Race/ethnicity is classified as non-Hispanic whites (reference group), non-Hispanic blacks, Hispanics, and other race (e.g., Asians, Native Americans, etc.). Other relevant labor market characteristics consist of actual work experience, actual work experience squared, residence in the South, urban residence, full-/part-time work status, and local unemployment rate (more than or equal to 6% or not) at the time hourly wages were measured. Actual work experience is calculated using the work history data from the NLSY79.

Method

This paper uses a quantile regression method to address the role of cognitive and noncognitive skills in between- and within-education group wage inequality. A conventional approach to examining the economic return to education is to employ ordinary least squares (OLS) models, which regress individual wages on educational attainment and observable group attributes. With the conditional means as a measure of central tendency, OLS provides a parsimonious description of the association between the explanatory and dependent variables. It facilitates accounting for between-education group variations in different dimensions, and then interprets the distribution of the wage residuals as capturing within-education group wage inequality.

However, OLS has at least two drawbacks inherent to the conditional mean models (Hao and Naiman 2007). First of all, they are not readily extended to detecting the relationship between the explanatory and dependent variables in non-central locations of the whole wage distribution. Given the natural interest of economic inequality and mobility in the poor (lower tail) and the rich (upper tail), an important question is whether and how educational attainment and other relevant explanatory variables—e.g., skill differences—have differential effect in various locations of the wage distribution. In this vein, the conditional mean models may be an inefficient way to delineate a comprehensive picture of between-education group wage inequality. Moreover, the emphasis of OLS on central location tends to pay less attention to the shape of the wage distribution, because it assumes that the effects of the explanatory variables on the conditional mean of wages are constant across the whole distribution. Inspection of changes in the shape of the wage distribution due to the explanatory variables, however, could provide a basis for investigating within-education group wage inequality. To the extent that the wage effect of educational attainment changes as other relevant explanatory variables have varying effects along the wage distribution, we can specify how the determinants of wage differentials operate within group with similar education. These weaknesses seem to make it difficult for traditional regression analysis such as OLS to directly link between- and within-education group wage inequality.

As an alternative analytic approach, quantile regression estimates the potential differential effects of the explanatory variables on various quantiles in the conditional wage distribution (Buchinsky 1998; Hao and Naiman 2007; Koenker and Bassett 1978; Koenker and Hallock 2001). *Quantile* refers to a generalized case of quartile, quintile, decile, and percentile, so it can be specified at any points of a distribution. For example, the 10^{th} quantile indicates 10% of a population lies below that quantile. Because of its distribution-based approach, quantile regression is capable of simultaneously describing both between- and within-education group wage inequality by evaluating the effect of the explanatory variables both within each conditional quantile (e.g., the median) and between conditional quantiles (e.g., the 10^{th} vs. 90^{th} quantiles) (Buchinsky 1994). Let $i=1, \ldots, n$ be a sample of individuals, y_i is the log transformed wage for individual i, and x_i is a $K \times 1$ vector of covariates. Quantile regression can be written as

$$y_i = x_i\,\beta_\theta + u_{\theta i}, \quad \text{Quant}_\theta(y_i|x_i) = x_i\,\beta_\theta, \quad \text{where } 0 < \theta < 1. \tag{1}$$

In Equation 1, θ points to the cumulative proportion of the sample, $\text{Quant}_\theta(y_i|x_i)$ denotes the θ^{th} quantile of y_i, conditional on the vector of the explanatory variables x_i, and $u_{\theta i}$ is the error term at the θ^{th} quantile that has zero expectation. The quantile regression estimator is analogous to that of OLS but has one different feature. While the least-squares estimator is obtained by taking the values of the parameters of the explanatory variables that minimize the sum of squared residuals, the quantile regression estimator solves for the parameters by minimizing a weighted sum of absolute residuals.[15] This can be written as

$$\min_{\beta \in R^k}\left\{\sum_{i:y_i \geq x_i\beta_\theta} \theta\,|\,y_i - x_i\beta_\theta\,| + \sum_{i:y_i < x_i\beta_\theta}(1-\theta)\,|\,y_i - x_i\beta_\theta\,|\right\} = \min_{\beta \in R^k}\left\{\sum_i \rho_\theta(y_i - x_i\beta_\theta)\right\}, \tag{2}$$

where $\rho_\theta(\lambda) = \theta\lambda$ if $\lambda \geq 0$ or $\rho_\theta(\lambda) = (\theta - 1)\lambda$ if $\lambda < 0$. From Equation 2, we can see that different weights are assigned to positive and negative residuals. For example, when estimating the parameters of the explanatory variables for the 90$^{\text{th}}$ quantile regression, 10% of observations with positive residuals are given a weight of .9 and the rest with negative residuals are given a weight of .1.

It should be noted that quantile regression should not be understood as segmenting the outcome variable into subsets and running the OLS regressions fitting on these subsets, which is doomed to distortion due to sample selection (Koenker and Hallock 2001). Since the quantile regression function is the weighted sum of absolute residuals, estimation for each location is based on the whole sample, not the subsets of the sample. Quantile regression estimates can be interpreted in the same fashion as in OLS estimates. In the following analyses, OLS regression results are also presented for the purpose of comparison.

[15] Minimizing the sum of absolute residuals is equivalent to solving a linear programming problem and in the analysis, parameter estimates are generated using the STATA *sqreg* command. It calculates standard errors for the explanatory variables using the bootstrap method recommended by Buchinsky (1998).

RESULTS

Descriptive Results

Table 2.1 presents the descriptive statistics of hourly wages and covariates by highest education completed in the NLSY79 2002 sample. Since the college wage premium is of main interest in this study, I focus on high school and college graduates. Consistent with the wage inequality literature, college graduates earn higher wages and possess higher cognitive and noncognitive skills, compared to high school graduates. Also, they are more likely than high school graduates to be white, come from two biological-parent families, and have parents with college education. During adolescence, college graduates were more likely than high school graduates to have parents who were in professional or management occupations, have higher family income, have fewer siblings, not live in the South and live in urban areas. At the time the information on hourly wages were collected, college graduates had more labor market experience, were less likely to live in the South, more likely to live in urban areas, and worked in local areas of which unemployment rate was lower, compared to high school graduates.

<< Table 2.1 about here>>

With respect to the relationship between the economic return to education and cognitive and noncognitive skills, the descriptive results in Table 2.1 reveal several patterns that warrant further investigation. First, as shown in the variance of hourly wages within each education group, wage inequality among college graduates is larger than that among high school graduates. Second, the gap in cognitive skills between high school and college graduates is wider than that in noncognitive skills. Third, the variance of noncognitive skills is larger among college graduates than among high school graduates, which is opposite to that of cognitive skills. These patterns suggest that although cognitive and noncognitive skills may influence both the college wage premium and wage inequality within college graduates, the role that each skill plays in these two kinds of wage inequality may be different. While difference in cognitive skills appears to be more concerned with wage inequality between high school and college graduates, difference in noncognitive skills appears to be more related to wage inequality within college graduates. I scrutinize these descriptive findings in the following multivariate results section by employing the quantile regression method.

The Role of Cognitive and Noncognitive Skills in the College Wage Premium and Wage Inequality within College Graduates

Table 2.2 reports quantile regression estimates of the wage premium for college education and cognitive and noncognitive skills using the 2002 sample. Note that each model includes other educational attainment dummies and all control variables, so that a stricter version of the college wage premium is estimated throughout the analysis.[16] As seen in Panel A (Model A with no cognitive and noncognitive skills as the explanatory variables), the college premium is highly significant across the whole wage distribution and larger at the upper portion of the wage distribution. While the OLS estimate indicates that on average, college graduates earn approximately 53 percent (=$e^{.423}$) more hourly wages than high school graduates, the quantile regression estimates show that the college premium is associated with a 33 percent increase in hourly wages at the 10^{th} percentile and with a 63 percent increase at the 90^{th} percentile. This differential effect of the college premium results in a significant wage inequality within college graduates. The second row of Panel A reports that the college wage premium coefficients statistically differ between the lower and upper tails in the wage distribution.

<< Table 2.2 about here >>

In Panels B and C, cognitive (Model B) and noncognitive (Model C) skills are introduced respectively. Both panels show that cognitive and noncognitive skills have a positive impact

[16] A common human capital model predicting wages is based on the Mincer equation, where log wages are modeled as the sum of a linear function of years of education and a quadratic function of years of potential labor market experience.

on wages in almost all locations of the distribution and are generally larger at the upper portion of the wage distribution. For example, one standard deviation increase in cognitive (noncognitive) skills is associated with a 6 percent (4 percent) increase in hourly wages at the 10^{th} percentile, whereas it is associated with a 9 percent (7 percent) increase at the 90^{th} percentile. As a result, the magnitudes of the college premium coefficients reduce by about 8 to 21 percent (0.4 to 17 percent) across the wage distribution due to taking cognitive (noncognitive) skills into account, although they remain significant. In addition, compared to Model A, Models B and C indicate that workers' differences in cognitive and noncognitive skills decrease differences in the college premium coefficients between the lower and upper tails in the wage distribution. While the 95^{th} vs. 5^{th} percentile difference is statistically significant in both models (a 12 percent and 17 percent reduction, respectively), the 90^{th} vs. 10^{th} percentile difference is barely significant when controlling for cognitive skills (a 24 percent reduction) and becomes insignificant when controlling for noncognitive skills (a 40 percent reduction).

Panel D presents the quantile regression results from the full model (Model D) that includes both cognitive and noncognitive skills. Some of the noncognitive skill coefficients become insignificant and yet the effect of each skill is generally consistent with that in Models B and C. In Model D, the college premium coefficients are still significant but their magnitudes reduce by about 14 to 31 percent across the wage distribution. The college premium is associated with a 28 percent increase in hourly wages at the 10^{th} percentile and with a 40 percent increase at the 90^{th} percentile. What is striking from these results is that compared to Model A, the 90^{th} vs. 10^{th} percentile difference in the college premium coefficients becomes insignificant after controlling for cognitive and noncognitive skills (a 53 percent reduction). In summary, these findings lend strong support to the premise of the economic return to education, captured by the college wage premium. It holds true even after controlling for cognitive and noncognitive skill differences and regardless of where in the wage distribution the college premium is estimated. Besides the robust wage effect of college education, however, the results suggest that cognitive and, to a lesser degree, noncognitive skills reduce not only between-education group wage inequality but also wage inequality within college graduates.[17] Also, as indicated in Models B to D, the effect of cognitive and noncognitive skills is more pronounced in the upper portion of the wage distribution. This implies that the polarization of the recent U.S. wage structure driven by the upper half wage inequality is due in part to the demand for cognitive and noncognitive skills among high wage jobs.

<< Table 2.3 about here >>

Obviously, these findings do not mean cognitive and noncognitive skills fully account for wage inequality within college graduates, because by definition, levels of hourly wages are much higher in the upper tail of the wage distribution. So using the quantile regression models in Table 2.2, I calculate the predicted hourly wages (in 2000 dollars) of college graduates at the 10th and 90th percentiles in the wage distribution, in order to gauge the extent to which cognitive and noncognitive skills explain wage inequality among college graduates.[18] Results appear in Table 2.3. Model A shows that the college premium difference between the lower and upper tails of the wage distribution is about 42 dollars when including only educational attainment and all control variables. As shown in Models B and C, introducing cognitive and noncognitive skills explains 10 percent and 7 percent of wage inequality among college graduates, respectively. Finally, Model D reports that the combined effect of cognitive and noncognitive skills accounts for 13 percent of wage inequality within college graduates. This effect size seems nontrivial but modest. Recall, however, that both

---

[17] In a supplemental analysis, I added respondents' log wages at the time of labor market entry to the quantile regression models. These wage change models provide a stricter test of the wage effect of cognitive and noncognitive skills, controlling for unobserved skill differences that are not captured in the models in Table 2.2. Results do not alter the findings reported here (available upon request from the author).

[18] See the notes in Table 2.3 for how to calculate the predicted hourly wages.

skill measures are constructed as an early predictor of workers' earnings and in general terms rather than in job-specific terms. It is possible, therefore, that the effect size is likely to increase if we have more comprehensive and timely approximate measures of cognitive and noncognitive skills. Although what other factors could explain the rest of wage dispersions among college graduates is beyond the scope of this paper, I conjecture that macro-level demand- and supply-side factors, which are treated as constant in the analysis, should have a substantial influence on wage inequality within college graduates. The recent literature has called into attention such factors as deindustrialization, organizational reconfiguration, SBTC, and the composition effect of educational expansion, for better understanding overall wage inequality (Autor, Katz, and Kearney 2006; Lemieux 2006; Morris and Western 1999; Neckerman and Torche 2007). In the conclusion section, I discuss other potential factors of wage inequality within college graduates.

The Contingency of the Wage Effect of Cognitive and Noncognitive Skills on the College Wage Premium

To further examine the ways in which the effect of cognitive and noncognitive skills on wage differentials is conditioned by the college premium, I estimate the interaction effects on wages between both skills and the college premium. Table 2.4 presents quantile regression estimates of these interaction effects. As seen in Panel A, when cognitive skills interact with the college premium, both their main and interaction effects are mostly concentrated at the locations above the median in the wage distribution with some exceptions at the 90th percentile for the main effect and at the 95th percentile for the interaction effect. On the other hand, Panel B shows that while noncognitive skills have a main effect at the middle portion of the wage distribution, they have a strong interaction effect at the upper tails. These results suggest that cognitive skills come into play for both between-education group wage inequality and wage inequality among college graduates, whereas noncognitive skills play a more important role in wage inequality among college graduates than in between-education group wage inequality.

<< Table 2.4 about here >>

In Panel C, the effect of both skills is estimated conditional on the college wage premium to examine if the results above hold true. In general, cognitive skills are found to have a main effect on wages in a similar fashion as in Panel A but their effect is not dependent on the college premium. Unlike the OLS estimate, there is no statistically significant interaction effect of cognitive skills along the wage distribution. In contrast, noncognitive skills do not have the strong main effect but their effect is highly contingent on the college premium at the upper tail of the wage distribution. The OLS estimate may be misleading in this regard, because of its failure to detect this interaction effect of noncognitive skills in specific segments of the wage distribution. The quantile regression interaction models, therefore, provide a more nuanced picture relative to the OLS interaction models. As reported in Panel C, the OLS result indicates that although both cognitive and noncognitive skills contribute to a reduction in between-education group wage inequality, it is only cognitive skills that matter in wage inequality among college graduates. The quantile regression result, however, shows that the wage effect of cognitive skills runs more noticeably between high school and college graduates and that of noncognitive skills does so within college graduates.[19] These findings suggest a distinctive role of noncognitive skills in patterns of the economic return to education. This may be unexpected in the unobserved ability bias story, which tends to assume that unobserved skills, whether cognitive or noncognitive, operate in a similar fashion in explaining between- and within-education group wage inequality.

[19] It should be noted that this result does not mean the effect of cognitive (noncognitive) skills is irrelevant to explaining within- (between-) education group wage inequality. As discussed in the previous section, both skills contribute to reductions in the college wage premium and wage dispersions within college graduates. The result from the quantile regression interaction model simply confirms that the wage effect of each skill operates quite differently in explicating between- and within-education group wage inequality.

The Role of Aging and the Accumulation of Labor Market Experience in the Relation between Cognitive and Noncognitive Skills and the College Wage Premium

Finally, I address whether the contribution of cognitive and noncognitive skills in accounting for between- and within-education group wage inequality changes due to workers' aging and accumulation of labor market experiences. [20] The quantile regression models in Tables 2 and 3 are re-estimated for a series of biannual samples. Figure 2.1 plots a growth in the college wage premium—wage difference between high school and college graduates—among the sample cohort at the lower, the median, and the upper tails of the wage distribution from 1990 to 2002, before and after controlling for cognitive and noncognitive skills. At the 10th percentile, the effect of cognitive and noncognitive skills on the college wage premium is small. There is little increase in the college premium; taking these skills into account reduces it by 7 percent in 1990 and 19 percent in 2002. Changes in the college wage premium at the median display a similar pattern, which indicates that cognitive and noncognitive skills explain 16 percent of the wage difference between high school and college graduates in both 1990 and 2002. In the meantime, both skills play a more important role in the growth in the college wage premium at the 90th percentile as it increases steadily throughout the period covered in this analysis. In 1990, controlling for cognitive and noncognitive skills accounts for only 13 percent of the college premium, whereas in 2002, both skills reduce it by 29 percent. In this vein, the effect of skill differences on between-education group wage inequality tends to amplify as workers age and spend more time in the labor market and at the same time, their effect becomes stronger at the upper tail of the wage distribution.

[20] The analysis for this section is agnostic about whether and how much the results presented here are confounded by period effect. Note, however, that the related literature shows a stable increase in wage inequality during this period, though to a lesser degree than during the 1980s (Autor, Katz, and Kearney 2006; Lemieux 2006). Random- or fixed-effects models could be used for taking potential period effects into consideration, but they are yet to be implemented in the quantile regression framework. In the analysis, all available workers are included in each of the biannual samples, so a substantial number of workers are not present across all waves.

<< Figure 2.1 about here >>

Figure 2.2 depicts the effect of cognitive and noncognitive skills on wage inequality within college graduates in their course of approaching prime ages in the labor market. As wage inequality among college-educated workers increases, differences in both skills explain 5 percent of wage difference between the 10^{th} and 90^{th} percentiles of the wage distribution in 1990 and 13 percent in 2002. As college graduates accumulate more labor market experience, skill differences play the more salient role in wage dispersions among them. The finding in Figure 2.1 has an important consequence for this result. It suggests that the impact of cognitive and noncognitive skills on wage inequality within college graduates is derived from their marked impact on the upper-half wage inequality. This finding suggests that skill differences could function as one of the determinants of the polarization of the recent U.S. wage structure among prime-aged workers.

<< Figure 2.2 about here >>

Taken together, the analysis in this section helps to elaborate the conclusion of Dunifon and Duncan (1998) and Dunifon, Duncan, and Brooks-Gunn (2001) that workers' early skill predictors have a significant effect on wages at much later ages. The long-run effect of cognitive and noncognitive skills is manifest in both between- and within-education group wage inequality. There are several scenarios that have the potential to disentangle this pattern. First, employers' assessment of workers' general skills needs a substantial amount of time. Second, as workers reach their prime ages in the labor market, cognitive and noncognitive skills formed during the early life stage synergize requirements for high ranking jobs, such as a sense of organization, efficiency, and a future orientation. Third, workers with higher levels of cognitive and noncognitive skills are more likely to invest on-the-job-training, which would result in higher wages at later ages. Fourth, they perform better in wage bargaining in job mobility processes. While each scenario remains to be empirically tested, the results reported here clearly show that the wage effect of early skill differences is positively associated with workers' aging and accumulation of labor market experience.

DISCUSSION AND CONCLUSION

Scholarship on labor market inequality in the United States has specified the rising economic return to education as a major source of changes in and levels of wage inequality since 1980s. Some researchers have investigated the potential of unobserved skill bias, which is conventionally captured by cognitive skills, to illuminate the education-labor market nexus. Still, this line of research has not resolved why between-education group wage inequality is not a satisfactory explanation for wage inequality, why the upper-half wage inequality contributes greater to overall wage inequality, and what sources other than cognitive skills drive within-education group wage inequality. This study is designed to provide a refinement of the explanation based on the economic return to education: 1) the role of noncognitive skills is assessed alongside cognitive skills by conceptualizing them as a form of cultural capital and a critical factor to cope with the principal-agent problem in the labor market; and 2) quantile regression models are employed to simultaneously evaluate the impact of cognitive and noncognitive skills on between- and within-education group wage inequality.

In summary, the quantile regression analysis provides four main findings. First, the economic return to education is by and in itself a robust explanation for changes in and levels of wage inequality. The college wage premium is highly significant, substantial in magnitude, and larger above the median of the wage distribution, even after holding constant cognitive and noncognitive skills as an early predictor of earnings as well as an extensive set of control variables. The college premium represents the effects of cognitive and noncognitive skills learned in schools, school quality, and educational credentials ("sheep-skin effect") among others. Second, while the above finding does not enhance our understanding of how the economic return to education involves within-education group wage inequality, this study finds that cognitive and noncognitive skills as pre-schooling and pre-labor market factors contribute significantly to reducing both wage differentials between high school and college graduates and wage inequality within college graduates. This result is mostly driven by the stronger impact of both skills at the upper portion of the wage distribution, leveling off the larger effect of the college premium above the median. In this vein, differences in cognitive and noncognitive skills function as an important factor of the polarization of the recent U.S. wage structure, because these skills are more valued among high wage jobs.

A third finding is that cognitive and noncognitive skills play quite different roles in between- and within-education group wage inequality. The results from the quantile regression interaction models show that when interacting with the college premium, cognitive skills are more influential in reducing wage differentials between high school and college graduates, whereas noncognitive skills are more pronounced in accounting for wage dispersions within college graduates, especially at the upper tail of the wage distribution. These results suggest that for college-educated workers at high-paying jobs, employers who are constantly faced with the principal-agent problem tend to reward a wage premium to them if they have high noncognitive skills. Given that these jobs allow for more autonomy, college-educated workers with a higher level of cultural capital (e.g., self-monitoring, internalized locus of control, and self-esteem) are more likely to respond positively to the employer-setting incentive structure. Lastly, cognitive and noncognitive skills as an early predictor of earnings have the long-term effect on between- and within-education group wage inequality. Their effect becomes more salient on both high school-college wage differentials—especially at the upper tail of the wage distribution—and wage inequality within college graduates, as workers approach their prime ages in the labor market. It suggests a cumulative nature of the association between the wage effect of early skill differences and workers' time spent in the labor market, which is consistent with Heckman and his colleagues' argument that "skill begets skill" (e.g., Carneiro and Heckman 2003).

These findings have substantive and methodological implications for the literature on the economic return to education. First of all, this study provides empirical evidence that skills are not unidimensional. Although most research regards skills as cognitive, noncognitive skills make distinctive contribution to explaining between- and within-education group wage inequality. In addition, given that cognitive and noncognitive skills are unevenly distributed among children with different levels of parental socioeconomic status and parenting practice, this study gives credence to research that emphasizes the family as an important institutional actor responsible for wage inequality. Morris and Western (1999) and Neckerman and Torche (2007) both recognize that more research should be conducted to explore the impacts of economic and political institutions on labor market inequality. The findings suggest that family-level response to the demand for skills also should be placed among these institutional factors. With regard to methodological issues, this study demonstrates that students of wage inequality can find the quantile regression models useful. In cases where the determinants of wages are theoretically thought to have differential effects across the wage distribution, the quantile regression method offers an alternative analytic approach to traditional regression methods.

Based on the results presented here, this study identifies several research foci that deserve attention. First, a closer examination of early skill formation is needed to uncover the process by which family background is associated with socioeconomic outcomes. A better understanding of the parent-child relationship and the gene-environment interactions is a high priority. In doing so, developing more elaborated measures of cognitive and noncognitive skills should be given another priority. Even if this study constructs the pre-schooling, pre-market measures of these skills to alleviate the measurement error problem, they are still limited in that the NLSY79 only gathered information on AFQT and personality traits. Assessing multiple indices of aptitude and behavioral traits (e.g., child problem behaviors) would increase the reliability of the measures of cognitive and noncognitive skills.

A second research focus is on specifying the mechanisms by which workers' early skill differences relate to their labor market performance under institutional constraints. Taking the finding of the direct wage effect of cognitive and noncognitive skills as a starting point, one could examine the linkages between early skill differences and more proximate indicators of wage determination that are heavily influenced by structure of work and occupation. Among such indicators are job-specific skills required at the firm and occupation levels and job mobility.

Third, sources other than early skill differences of within-education group wage inequality need to be addressed. This study identifies the contribution of cognitive and noncognitive skills to explaining wage inequality within college graduates, but large portion is still left unexplained. Overeducation, school quality, and the different field of study could drive larger wage dispersions among college graduates (Martins and Pereira 2004). Those factors can be important to the extent that the highly educated end up getting low wage jobs because there are so many qualified people for high wage jobs, colleges are more stratified in terms of school quality and reputation, and individuals' choice of the field of study matters in wage determination. Studies of within-education group wage inequality should take seriously these factors along with skill differences. Lastly, but not least, more research should be done to link the wage effect of skill differences with wage inequality by gender and race/ethnicity. Supplemental analyses (available upon request from the author) find that skill differences are stronger for men than women and to a lesser degree, for whites than minority groups. Therefore, future research needs to examine whether and how potential institutional arrangements (e.g., female wage penalty or statistical discrimination by race/ethnicity) may suppress the role of skill differences in wage inequality.

In conclusion, a careful consideration of differences in cognitive and noncognitive skills enriches research on the economic return to education by extending its premise to both between- and within-education group wage inequality in contemporary U.S. society. This paper also joins an emerging literature that calls more attention to the relationship between intergenerational mobility and socioeconomic inequality (Breier 1995). Despite an intuitive connection between the two, studies of mobility and inequality have been strikingly detached to each other (Hout 2004). I expect that sociological efforts to link these two will benefit from incorporating early skill formation process and its lasting impact on adult life outcomes.

Table 2.1. Descriptive Statistics by Educational Attainment, NLSY79 2002 Sample (unweighted)

		Highest Education Completed			
Variable	All	< High School	High School	Some College	College and More
Dependent Variable					
Log Hourly Wages	2.95	2.69	2.80	2.97	3.33
	(0.58)	(0.50)	(0.52)	(0.52)	(0.59)
Gap between Groups[a]		0.11	—	0.18	0.54
Variance within Each Group		0.25	0.27	0.27	0.35
Skill Measures					
Cognitive Skills	0.06	-0.42	-0.31	0.09	0.91
	(1.02)	(0.79)	(0.92)	(0.88)	(0.86)
Gap between Groups[a]		0.11	—	0.40	1.22
Variance within Each Group		0.63	0.84	0.88	0.74
Noncognitive Skills	0.01	-0.18	-0.19	0.11	0.33
	(0.99)	(0.94)	(0.91)	(1.04)	(0.99)
Gap between Groups[a]		-0.01	—	0.30	0.52
Variance within Each Group		0.88	0.84	1.08	0.99
Control Variables					
Sex (Male=1)	0.51	0.59	0.53	0.43	0.51
White (reference)	0.43	0.31	0.41	0.37	0.60
Black	0.29	0.33	0.32	0.35	0.18
Hispanic	0.18	0.25	0.18	0.22	0.12
Other Race	0.09	0.11	0.10	0.07	0.11
Two Biological Parent Family at Age 14 (reference)	0.68	0.53	0.67	0.66	0.80
Step Family at Age 14	0.09	0.14	0.10	0.07	0.05
Single Mother Family at Age 14	0.18	0.24	0.18	0.20	0.12
Other Family Structure at Age 14	0.06	0.09	0.06	0.07	0.03
Parental Education at Age 14: Less Than High School	0.32	0.54	0.38	0.30	0.11
Parental Education at Age 14: High School (reference)	0.41	0.37	0.48	0.40	0.32
Parental Education at Age 14: Some College	0.12	0.07	0.08	0.16	0.17
Parental Education at Age 14: College or More	0.15	0.03	0.05	0.13	0.40
Parental Occupation in 1979: Professional/Management	0.19	0.10	0.10	0.19	0.39
Family Income in 1979/$1000	17.12	12.79	15.87	16.24	22.81
	(11.13)	(7.70)	(9.80)	(9.61)	(13.89)
Number of Siblings in 1979	3.62	4.52	3.91	3.40	2.78
	(2.57)	(2.93)	(2.58)	(2.57)	(1.98)
Residence in the South at Age 14	0.38	0.44	0.39	0.40	0.31
Urban Residence at Age 14	0.79	0.82	0.76	0.80	0.80
Full-/Part-Time Work Status	0.84	0.78	0.85	0.85	0.87
Actual Work Experience	17.68	15.46	17.71	17.89	18.86
	(4.54)	(5.22)	(4.68)	(4.43)	(3.24)
Residence in the South	0.42	0.45	0.42	0.44	0.36
Urban Residence	0.75	0.77	0.69	0.77	0.80
Unemployment Rate ≥ 6%	0.53	0.54	0.55	0.53	0.49

Notes: *N*=2,527; Standard deviations for interval variables in parentheses; Missing indicators for parental occupation and family income are included in the analysis but not shown.

[a] indicates difference between each education group and high school graduates.

Table 2.2. Parameter Estimates from Quantile Regression Models for Log Hourly Wages, NLSY79 2002 Sample

	Quantile Regression							OLS
	q.05	q.10	q.25	q.50	q.75	q.90	q.95	
A. College Premium								
College Premium	0.242 *	0.288 *	0.432 *	0.459 *	0.475 *	0.489 *	0.477 *	0.423 *
	(0.059)	(0.046)	(0.039)	(0.032)	(0.043)	(0.069)	(0.076)	(0.029)
Equality of College		q.95 vs. q.05:	0.235 *		q.90 vs. q.10:	0.201 *		
Premium Coefficients[a]			[0.020]			[0.011]		
B. College Premium + Cognitive Skills								
College Premium	0.198 *	0.234 *	0.397 *	0.397 *	0.414 *	0.387 *	0.404 *	0.369 *
	(0.068)	(0.047)	(0.042)	(0.035)	(0.051)	(0.066)	(0.082)	(0.031)
Cognitive Skills	0.032	0.058 *	0.043 *	0.076 *	0.060 *	0.085 *	0.101 *	0.060 *
	(0.028)	(0.021)	(0.017)	(0.015)	(0.021)	(0.026)	(0.031)	(0.013)
Between-Group[b]	0.044	0.054	0.036	0.062	0.060	0.102	0.073	0.054
% Explained	18.3	18.7	8.3	13.5	12.7	20.9	15.3	12.7
Equality of College		q.95 vs. q.05:	0.206 *		q.90 vs. q.10:	0.153 †		
Premium Coefficients[a]			[0.043]			[0.051]		
% Explained			12.3			24.1		
C. College Premium + Noncognitive Skills								
College Premium	0.235 *	0.287 *	0.406 *	0.450 *	0.433 *	0.408 *	0.431 *	0.403 *
	(0.063)	(0.042)	(0.035)	(0.033)	(0.042)	(0.069)	(0.081)	(0.029)
Noncognitive Skills	0.036 †	0.035 *	0.027 †	0.038 *	0.055 *	0.067 *	0.081 *	0.045 *
	(0.019)	(0.013)	(0.014)	(0.012)	(0.015)	(0.024)	(0.032)	(0.010)
Between-Group[b]	0.007	0.001	0.027	0.009	0.042	0.081	0.046	0.020
% Explained	2.9	0.4	6.2	2.0	8.9	16.5	9.6	4.7
Equality of College		q.95 vs. q.05:	0.196 *		q.90 vs. q.10:	0.121		
Premium Coefficients[a]			[0.049]			[0.117]		
% Explained			16.6			39.7		
D. College Premium + Cognitive and Noncognitive Skills								
College Premium	0.185 *	0.243 *	0.370 *	0.396 *	0.387 *	0.338 *	0.372 *	0.362 *
	(0.067)	(0.045)	(0.042)	(0.034)	(0.044)	(0.070)	(0.085)	(0.031)
Cognitive Skills	0.030	0.042 †	0.037 *	0.061 *	0.055 *	0.076 *	0.101 *	0.050 *
	(0.030)	(0.021)	(0.019)	(0.015)	(0.019)	(0.028)	(0.033)	(0.013)
Noncognitive Skills	0.032	0.030 *	0.025	0.025 *	0.046 *	0.054 *	0.039	0.037 *
	(0.021)	(0.015)	(0.015)	(0.012)	(0.017)	(0.024)	(0.034)	(0.010)
Between-Group[b]	0.058	0.045	0.062	0.063	0.088	0.151	0.105	0.061
% Explained	23.8	15.6	14.4	13.7	18.4	30.9	22.0	14.4
Equality of College		q.95 vs. q.05:	0.188 †		q.90 vs. q.10:	0.095		
Premium Coefficients[a]			[0.087]			[0.251]		
% Explained			20.0			52.9		

Notes: N=2,527; Bootstrap standard errors in parentheses (300 replications); p-values from F-statistics in brackets; Each model includes all control variables but not shown.

[a] tests statistical difference in the college wage premium coefficients between the lower and upper tails in the wage distribution.

[b] indicates differences in the college wage premium coefficients before and after controlling for skill differences across the wage distribution.

† <.10; * p < .05 (two-tailed tests).

Table 2.3. Percent Reduction in Wage Inequality within College Graduates by Skill Differences, NLSY79 2002 Sample

	Predicted Hourly Wages in 2000 Dollars			
	q.10	q.90	Diff.	% Explained
Model A. College Premium	14.567	56.559	41.992	
Model B. A + Cognitive Skills	13.741	51.630	37.889	10%
Model C. A + Noncognitive Skills	14.328	53.206	38.878	7%
Model D. A + Cognitive and Noncognitive Skills	13.877	50.387	36.509	13%

Notes: The predicted hourly wages are calculated by averaging each respondent's value on all covariates except for being a college graduate and cognitive and noncognitive skills. Both skills are set to their sample means. The resulting predicted log hourly wages are exponentiated to be shown in dollar terms; Each model includes all control variables.

Table 2.4. Parameter Estimates from Quantile Regression Interaction Models for Log Hourly Wages, NLSY79 2002 Sample

	Quantile Regression							OLS
	q.05	q.10	q.25	q.50	q.75	q.90	q.95	
A. College Premium X Cognitive Skills								
College Premium	0.188 *	0.243 *	0.345 *	0.315 *	0.317 *	0.304 *	0.232 *	0.314 *
	(0.079)	(0.050)	(0.048)	(0.035)	(0.049)	(0.071)	(0.096)	(0.035)
Cognitive Skills	0.011	0.035	0.037	0.065 *	0.058 *	0.069	0.104 †	0.041 *
	(0.040)	(0.024)	(0.024)	(0.020)	(0.026)	(0.043)	(0.056)	(0.019)
College Premium X	0.017	0.029	0.034	0.064 *	0.087 *	0.095 †	0.113	0.068 *
Cognitive Skills	(0.068)	(0.044)	(0.045)	(0.030)	(0.044)	(0.057)	(0.073)	(0.029)
B. College Premium X Noncognitive Skills								
College Premium	0.186 *	0.257 *	0.381 *	0.365 *	0.355 *	0.318 *	0.320 *	0.346 *
	(0.077)	(0.047)	(0.043)	(0.041)	(0.045)	(0.064)	(0.093)	(0.032)
Noncognitive Skills	0.015	0.031	0.022	0.037 †	0.058 *	0.036	-0.046	0.033 †
	(0.029)	(0.022)	(0.024)	(0.021)	(0.027)	(0.040)	(0.053)	(0.018)
College Premium X	0.038	0.020	0.008	0.029	0.053	0.144 *	0.183 *	0.053 *
Noncognitive Skills	(0.062)	(0.045)	(0.040)	(0.034)	(0.036)	(0.066)	(0.082)	(0.027)
C. College Premium X Cognitive Skills and College Premium X Noncognitive Skills								
College Premium	0.188 *	0.226 *	0.355 *	0.329 *	0.317 *	0.308 *	0.328 *	0.307 *
	(0.086)	(0.054)	(0.049)	(0.035)	(0.049)	(0.066)	(0.085)	(0.035)
Cognitive Skills	0.027	0.036	0.034	0.068 *	0.043	0.073 †	0.155 *	0.041 *
	(0.039)	(0.027)	(0.025)	(0.022)	(0.026)	(0.043)	(0.055)	(0.019)
Noncognitive Skills	0.023	0.027	0.023	0.038 †	0.053 †	0.038	-0.054	0.035 *
	(0.031)	(0.020)	(0.026)	(0.021)	(0.030)	(0.043)	(0.060)	(0.018)
College Premium X	0.019	0.042	0.035	0.053	0.076	0.039	-0.037	0.060 *
Cognitive Skills	(0.074)	(0.048)	(0.048)	(0.035)	(0.048)	(0.063)	(0.076)	(0.030)
College Premium X	0.025	0.022	0.015	0.016	0.048	0.131 *	0.196 *	0.044
Noncognitive Skills	(0.065)	(0.047)	(0.043)	(0.031)	(0.036)	(0.062)	(0.081)	(0.027)

Notes: N=2,527; Bootstrap standard errors in parentheses (300 replications); Models A and B also estimate the main effect of noncognitive and cognitive skills, respectively, but not shown; Each model includes all control variables but not shown.

† <.10; * p < .05 (two-tailed tests).

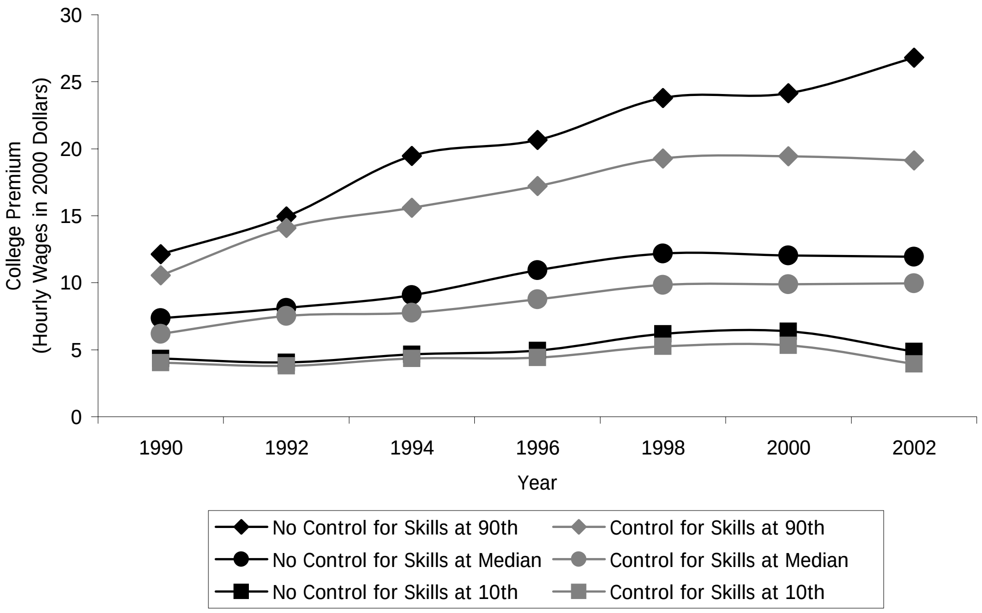


Figure 2.1. Evolution of the College Wage Premium, NLSY79 1990-2002 Samples

Notes: Figure 1 is based on quantile regression models in Table 2 that are estimated for each biannual sample; Each model includes all control variables.

Figure 2.2. Evolution of Wage Inequality within College Graduates, NLSY79 1990-2002 Samples

Notes: Figure 2 is based on quantile regression models in Table 3 that are estimated for each biannual sample; Each model includes all control variables.

CHAPTER 3

Educational Assortative Mating in Adolescence and Adulthood:
The Role of Cognitive and Noncognitive Skills

INTRODUCTION

Extensive research to date on socioeconomic inequality has suggested that education governs social mobility as it becomes a primary locus of human capital investment and the transmission of familial resources (Blau and Duncan 1967; Bourdieu 1984; Bowles and Gintis 2002a). In this light, scholarship on marriage has expressed a growing concern about socioeconomic inequality associated with increases in educational assortative mating in the United States (Fernández, Guner, and Knowles 2005; Kalmijn 1998; Mare 1991). The prevalence of the educational resemblance of spouses in union formation may indicate the rigidity of social stratification process. Previous research provides evidence for widened differentials in marriage rates by educational attainment and a rise in educational homogamy over time (Goldstein and Kenney 2001; Schwartz and Mare 2005). Trend analyses show that as who marries whom depends increasingly on educational attainment and earnings potential (Lichter et al. 1992; Sweeney 2002), educational assortative mating has increased over the past six decades even after taking demographic change in the education distribution into account. This line of research increases our knowledge about how education-based partner selection might relate to social stratification processes.

However, we still know strikingly little about how differentials in social, cultural, and economic resources lead to educational assortative mating and the ways in which it is intertwined with socioeconomic inequality. Although education has been offered as a major explanation for these linkages, it is problematic because education represents two contrasting forces—family background and a meritocratic turn in the U.S. society. While it is plausible that background factors have an indirect impact on one's education-based partner choice by their influence on his or her educational attainment, education also forges preferences for cultural similarity and provides opportunities for mate selection by functioning as local marriage markets (Kalmijn 1998). In order to differentiate these interrelated roles of education, we need to uncover unidentified sources that would clarify the relation between educational assortative mating and socioeconomic inequality. Furthermore, most research has yet to incorporate the *intimate* nature of the mating process into its empirical analysis. There has been little understanding of the role of personal traits in educational assortative mating. It is well recognized that both attitude/behavior-based fellowship and education-based mate selection are a form of homophily, but connecting these two has been rare (McPherson, Smith-Lovin, and Cook 2001). This lack of attention to attitudinal and behavioral characteristics may be due to the assumption that they are too idiosyncratic and weakly correlated with status to predict education-based partner choice (Kalmijn 1994). However, given that some of these traits (e.g., intelligence and self-esteem) are unevenly distributed across family socioeconomic status, it is necessary to take a closer examination of whether these personal traits and habits serve as an important link between family background and mate selection by education.

In this paper, I assess the role of cognitive and noncognitive skills in individual patterns of educational assortative mating to address its implications for socioeconomic inequality. Theoretically, this study builds on the literature on forms of capital to develop the concept of noncognitive skills as a form of cultural capital (Bourdieu 1984; Lareau 2002; Swidler 1986). Correlated but distinct from cognitive skills (e.g., intelligence), noncognitive skills refer to attitudinal and behavioral traits and habits, such as locus of control and self-esteem. An emerging literature on noncognitive skills has observed the potentials of both skills for individuals' educational attainment and labor market performance (Bowles and Gintis 2000;

Farkas 2003; Heckman, Stixrud, and Urzua 2006). It also points out that while noncognitive skills are less innate and more malleable than cognitive skills, formation of both skills mostly occurs during childhood and adolescence and is heavily influenced by family resources and parent-child relationship (Carneiro and Heckman 2003). Based on these findings, this study investigates whether and how the significant effects of cognitive and noncognitive skills on multiple life domains found in the prior research apply to individuals' propensity to educational assortative mating.

Analytically, this study improves upon the extant literature by examining the linkages between skill differences and men's and women's education-based partner choice in both adolescence and adulthood, using the data from the National Longitudinal Study of Adolescent Health (Add Health) and the National Longitudinal Survey of Youth 1979 (NLSY79). Whether skill differences contribute to educational assortative romantic relationships in adolescence is a critical question, because it signifies that as far as education-based mate selection is concerned, social closure may persist even at the early life stage. Employing multinomial logistic regression and multinomial discrete-time event history analyses, this study tests the role of skill differences in educational assortative mating among both adolescents and adults.

Searching for the association between cognitive and noncognitive skills and educational assortative mating will provide a valuable insight into its relation to socioeconomic inequality. Elaborated status attainment models show that socioeconomic outcomes are very much contingent on differentials in cultural as well as economic resources that individuals utilize, besides educational attainment (Bowles and Gintis 2002b; DiMaggio and Mohr 1985; Lamont and Lareau 1988). If these resources have direct and/or indirect impacts on mate selection by education, then this process reduces interactions between different status groups, resulting in an increasing social divide. Given the convertibility of cultural capital into economic and social capital and the potential for cultural similarity to be a basis of the mate selection process (Bourdieu 2002; Kalmijn 1994), educational assortative mating via cognitive and noncognitive skills can be a key to understanding socioeconomic inequality alongside conventional socioeconomic outcomes.

PREVIOUS RESEARCH ON EDUCATIONAL ASSORTATIVE MATING

Educational Homogamy in Adulthood

Studies of trends in educational assortative mating in the United States have consistently reported that the educational resemblance of spouses increased during the second half of the twentieth century (Mare 1991; Pencavel 1998; Schwartz and Mare 2005). Although sluggish from 1940 to 1960 and in the 1980s, educational homogamy has continued to rise through the 1990s into the twenty-first century. When divided into discrete educational attainment groups, not only did the odds of intermarriage between those with college education and those with less education decrease but also the odds of marrying up among those with very low education declined.

Explanations for increases in educational assortative mating put these trends in a larger context of the changing patterns of marital behaviors, delving into changes in opportunities and preferences for educational homogamy. There are two interpretations of marriage patterns, based on educational expansion, rising female labor force participation, and changes in gender roles in households. Becker (1981)'s economic theory of marriage proposes the women's economic independence hypothesis, arguing that these social changes have been in favor of raising women's earnings potential, which led them to reconsider gains to marriage. As an ethic of expressive individualism—to follow Bellah and colleagues' (1985) term—became more important for decisions on life events, union formation has been depicted as the "retreat from marriage" or the so-called deinstitutionalization of marriage (Bumpass, Sweet, and Cherlin 1991; Cherlin 2004).

The other interpretation views the same factors described above as influencing the timing of marriage rather than marriage rates. It argues that young adults' increased education and women's increased labor force participation, both formulated as supporting the economic independence hypothesis, may result in marriage delayed but not marriage forgone, suggesting that these changes do not necessarily signify the considerable rise in nonmarriage (Oppenheimer 1988; Oppenheimer, Blossfeld, and Wackerow 1995). Moreover, the literature on the relationship between education and marriage does not parallel the economic independence hypothesis. Empirical findings are mixed at best: while some find a negative effect of educational attainment on marriage, recent evidence shows that the effect on marriage formation of educational attainment as a reliable proxy for women's economic independence has been increasingly positive over time, net of marriage timing and school enrollment (Qian and Preston 1993; Thornton, Axinn, and Teachman 1995).[1]

In general, the context of the association between education and marriage pinpoints patterns of educational assortative mating: It does not simply imply that the more educated have a higher propensity to marry but also that education functions as an important constraint to opportunities and preferences for mate selection. The aggregate-level analysis of educational assortative mating has focused on 1) a marriage market functioning of education and 2) individual preferences for educational similarity. First, marriage markets primarily refer to opportunity, or availability, given to marriage candidates. The expansion of education suggests a gaining influence of educational institutions as marriage markets, where

[1] These mixed findings are related to the level of analysis of marriage. For example, macro-level analyses appear to support the economic independence hypothesis, stressing the negative association between female labor force participation and marriage rates. In contrast, micro-level analyses show the positive association between the two. The use of longitudinal data gives more credence to the micro-level analyses in the sense that cross-sectional data that most macro-level analyses rely on are fragile to reverse causality.

young people are more homogeneous as levels of educational attainment increase and thereby more likely to marry homogamously (Schwartz and Mare 2005).[2]

Second, on the preference side, these structural constraints point to a growing social distance between individuals by educational attainment. To the extent that education signals future socioeconomic success and shared cultural taste that provide a basis for mate selection, marriage likely occurs according to the differentiation of education groups (Kalmijn 1998; Rosenfeld 2005; Schoen and Weinick 1993). Moreover, research has shown that the opportunity side of education does not fully account for patterns of education-based marital sorting, which leaves much room for considering marriage candidates' preference for educational resemblance (Blackwell 1998; Kalmijn 1998; Stevens 1991).

Although this body of literature identifies that education-based spouse selection is due in large part to individuals' propensity to homophily, however, it is inconclusive as to how social, cultural, and economic resources are associated with education-based marital homophily. Individuals' educational attainment ought to function as a link between the two, but it is difficult to isolate this aspect of education from its own role in formulating one's preference for mate selection and serving as marriage markets. In addition, despite the potential connection between marriage sorting by education and socioeconomic inequality, few identifiable mechanisms have developed into testable hypotheses.[3] In this study, I

[2] In a similar vein, local marriage markets might have an impact on educational homogamy in terms of their socioeconomic compositions. It is found that residents in educationally unfavorable local marriage markets are less likely to marry up (Lewis and Oppenheimer 2000).

[3] There may be several potential mechanisms by which to link educational assortative mating and socioeconomic inequality. One is to analyze this link at the aggregate level (Fernández, Guner, and Knowles 2005). Since this study uses survey data, this type of analysis is not immediately available. The other is, as some researchers suggested (Kalmijn 1994; Mare 1991), to examine how the current-level of educational assortative mating affects future generations' socioeconomic outcomes, but it can be made possible only if data on next generations are collected. It should be noted that my interest is in specifying the link between patterns of educational assortative mating and socioeconomic inequality when the former is an outcome. In complex causal chains, the opposite direction is also plausible depending on what causal direction is of interest.

explore a potential mechanism by which familial resources have a differential impact on individual patterns of education-based spouse choice via cognitive and noncognitive skills.

Educational Assortative Romantic Relationships in Adolescence

Unlike patterns of educational homogamy in adulthood, research on educational similarity in adolescent romantic relationships has been scant. Most studies concentrate on how adolescent romantic relationships impact academic and health outcomes during adolescence and mating behaviors during adulthood, rather than on how these relationships are formed in the first place (Joyner and Udry 2000; Raley, Crissey, and Muller 2007). This tendency in the literature appears to reflect the understanding of the strong social pressure on adolescents' dating behavior and the sometimes rocky nature of adolescence. However, a recent study finds that romantic relationships are common among the recent cohort of adolescents and the duration of these relationships are remarkably long (Carver, Joyner, and Udry 2003). It is reported that by age 18, 69 percent of boys and 76 percent of girls have had a romantic relationship in the last 18 months and half of the romantic relationships have been ongoing for at least 21 months among adolescents at age 16 and older.

Given the prevalence of romantic relationships during adolescence, research on adolescent friendship formation provides an important clue to address patterns of educational assortative partner choice in adolescence. Cohen (1977) and Kandel (1978) show that adolescent friendships are homophilous in terms of school achievement and educational aspirations.[4] They find that this educational homophily results more from selection into friendships with similar others than from peer influence. If this pattern is also found in adolescent romantic relationships, it would signify evidence for social segregation at this early stage in the life course. However, as implied in the interrelated roles of education in

[4] Wang, Kao, and Joyner (2006) also demonstrate the homophilic characteristic of adolescent romantic relationships with a focus on racial similarity. They find that between-racial group romantic relationships are less stable than within-racial group romantic relationships.

educational homogamy in adulthood, academic achievement also represents the influences of family background and schooling at the same time. In this respect, I expect that the role of cognitive and noncognitive skills in academic-performance-based partner choice in adolescence will shed an additional light on the relationship between educational assortative mating and socioeconomic inequality.

COGNITIVE AND NONCOGNITIVE SKILLS, EDUCATIONAL ASSORTATIVE MATING, AND SOCIOECONOMIC INEQUALITY

Effects of Cognitive and Noncognitive Skills on Socioeconomic Outcomes[5]

Cognitive skills pertain to a general intelligence, or the "*g*" factor, whereas noncognitive skills point to enduring dispositions that are not captured by cognitive skills.[6] These personal traits and habits include perseverance, self-confidence, sociability, emotional stability, interpersonal skills, and future orientation. While both cognitive and noncognitive skills are highly dependent on changes in social conditions, noncognitive skills are more likely to be formed during childhood and stabilize during adolescence, compared to cognitive skills that are likely to be shaped during early childhood (Bowles and Gintis 2000; Heckman and Rubinstein 2001).[7]

The notion of noncognitive skills can be better understood as a form of cultural capital. Its important aspect is found in Swidler (1986)'s extension of Bourdieu's (1977) concept of

---

[5] Since individual skill differences have been represented mostly by cognitive skills in the literature, this section focuses more on the relation of noncogntive skills to socioeconomic outcomes. For a review of the role of cognitive skills, see Farkas (2003).

[6] In the literature, noncognitive skills are used interchangeably with soft skills or socioemotional skills. The term "noncognitive skills" has been adopted to distinguish soft skills from hard skills, although there is no doubt that soft skills also involve cognitive processes (Farkas 2003; Kuhn and Weinberger 2005). The term "noncognitive skills" is used throughout this paper mainly because this kind of skill has not been closely addressed in most research on socioeconomic inequality, which primarily regards ability as cognitive skills.

[7] According to Carneiro and Heckman (2003), cognitive skills are fairly stable after age 8, and noncognitive skills can be improved until the late teenage years.

habitus. She considers culture as a “tool kit” that constructs “strategies of action”: “One can hardly pursue success in a world where the accepted skills, styles, and informal know-how are unfamiliar. One does better to look for a line of action for which one already has the cultural equipment” (Swidler 1986:275).[8] Another aspect of cultural capital is its convertibility into economic and social capital and *vice versa* (Bourdieu 1984; Bourdieu and Wacquant 1992). As education takes a leading role in social stratification system, cultural capital can be utilized in accumulation of human capital in ways to be rewarded and reinforced in schools and workplaces (Bowles and Gintis 2002a; Farkas et al. 1990; Rosenbaum 2001). Alongside cognitive skills, therefore, possession or lack of noncognitive skills influences individuals’ socioeconomic outcomes. In this respect, the conceptualization of noncognitive skills in a form of cultural capital reveals the importance of family background and parent-child relationship. As Farkas (2003) pointed out, cognitive and noncognitive skills formed in disadvantaged families may be essential to coping with poor economic situations but not be successful in schools and labor markets. Social, cultural, and economic resources available in the family are crucial in shaping and activating individuals’ cognitive and noncognitive skills.

With recognition of the role of familial resources in skill formation, a growing body of literature has provided rich evidence for the effects of cognitive and noncognitive skills on multiple domains of adolescent and adult outcomes. Heckman, Stixrud, and Urzua (2006) report that noncognitive skills have a strong impact on schooling decisions, especially on graduating from a four year college. Duckworth and Seligman (2005) show that self-discipline outweighs IQ in predicting adolescent academic performance. With respect to labor market outcomes, Jencks et al. (1979) present that noncognitive skills, such as perseverance, industriousness, and leadership, have a positive effect on wages even after

[8] DiMaggio (1982) defines cultural capital more formally as interest in and experience with prestigious cultural resources. While highlighting an important dimension of cultural capital, this definition is not as directly concerned with personal traits and habits as is Swidler’s.

controlling for a great number of human capital variables. Bowles, Gintis, and Osborne (2001) find that individuals' noncognitive skills, which they call "incentive-enhancing preferences," are as much rewarded in the labor market as cognitive skills. In an analysis of the General Educational Development (GED), Heckman and Rubinstein (2001) demonstrate that GED holders have higher cognitive skills than high school dropouts but earn lower wages than high school graduates, since they have lowest noncognitive skills among education groups. Rosembaum (2001) finds that leadership measured in high school strongly predicts earnings 10 years later, coupled with cognitive skills and other noncognitive skills.

The Role of Cognitive and Noncognitive Skills in Educational Assortative Mating

Whether cognitive and noncognitive skills favored in schools and labor markets are also valued in the education-based mate selection processes is an empirical question. The findings described above imply that individuals with high cognitive skills and socially valued noncognitive skills are more likely to date the academically achieved and to marry the more educated, all other things equal. Since both skills are formed and activated at the early life stage when educational attainment is a primary goal of individual development, and they are an early predictor of individuals' socioeconomic outcomes, skill differences are likely to lead to educational assortative mating. In addition, inclusion of cognitive and noncognitive skills helps mitigate the individual-level unobserved heterogeneity problem that tends to exacerbate analysis of the determinants of educational assortative mating. If mate selection essentially entails a matching process, which is characterized as forming intimate, personal relationships, one should consider measures of attitudinal and behavioral traits along with other sociodemographic characteristics to discern the sources of homophily.

It should be noted that the relationship between skill differences and individual patterns of educational resemblance of partners is likely to be constrained and mediated by several factors. First, given the salience of parental control during adolescence, parent-child relationship affects skill formation and at the same time constrains educational assortative romantic relationships among adolescents (Wang, Kao, Joyner 2006). School characteristics are considered another constraining factor because schools provide a structural environment for the absolute majority of adolescents. Students' other characteristics regarding academic performance and sexual behavior might mediate the correlation between skill differences and partner sorting by academic achievement. Second, current living arrangement and local marriage market characteristics condition individual patterns of education-based spouse choice (Goldscheider and Waite 1986; Lichter et al. 1992; Lewis and Oppenheimer 2000). Economic potential, such as education, earnings, and employment, should mediate the effect of skill differences, because both cognitive and noncognitive skills have positive influence on adult socioeconomic attainment. Finally, and most importantly, gender difference is likely to moderate manifestations of all of these factors. Men and women may perceive other gender's skill levels differently. So it is possible that despite some convergence in men's and women's partner choice behavior, there is a difference by gender in the association between skill differences and education-based mate selection. Thus, I conduct separate analyses by gender of the potential mechanism in which family background is associated with education-based partner choice via skill differences, taking the constraining and mediating factors into account.

Skill Differences as a Link between Educational Assortative Mating and Socioeconomic Inequality

Identification of this proposed mechanism has direct relevance to socioeconomic inequality. The strength of educational homogamy indicates the "openness" or "closure" of society, which is analogous to the relationship between the occupational similarity of fathers and sons (Lipset and Bendix 1959). Since the major portion of families and households is still made up by marriages, the degree of the educational resemblance of spouses would represent between-family inequality. Kalmijn (1991a, 1991b), therefore, argues that the prevalence of educational assortative mating indicates the rigidity of stratification system as a paradoxical result of a meritocratic character of American society. As principles of social organization were transformed from ascription-oriented to achievement-oriented, the role of "third parties," such as families, churches, and the state, diminishes in individuals' mate selection.

However, corresponding to these contextual changes, there may be family-level adjustments in ways to put more or less emphasis on cultural capital alongside economic capital (Blackwell 1998; Bowles and Gintis 2002b; DiMaggio and Mohr 1985). Indeed, elaborated status attainment models have shown that intergenerational mobility is less fluid than the earlier ones indicated (Corcoran 1995, Rytina 1992; Zimmerman 1992). Transmission of familial resources by past generation may affect current generation's education-based spouse selection not just via education but also via cognitive and noncognitive skills, which suggests a more nuanced pattern of intergenerational mobility. Furthermore, the extent to which skill differences affect adolescent partner choice by academic achievement provides another indicator of the role of educational assortative mating in socioeconomic inequality. If adolescents with higher levels of cognitive and noncognitive skills are more likely to date the academically achieved, this type of social segregation among adolescents means that social closure is present even at the early life stage in spite of the volatile nature of adolescence and large potential pool of diverse mates.

DATA, MEASURES, AND METHODS

Data

This study uses data from the National Longitudinal Study of Adolescent Health (Add Health) and the National Longitudinal Survey of Youth 1979 (NLSY79). Each data source is analyzed to examine educational assortative romantic relationship in adolescence and education-based spouse sorting in adulthood, respectively. Add Health is a nationally representative, school-based, longitudinal study of adolescents in grade 7 to 12 in 1994-1995 (Harris et al. 2003). The Add Health study employs a school-based design to select a stratified sample of 80 high schools with selection proportional to size. A feeder school per each high school was selected as well with probability proportional to its student contribution to the high school. Therefore, the school-based sample has a pair of schools in each of 80 communities. An in-school questionnaire was administered to more than 90,000 adolescents who attended each selected school on a particular day during the period of September 1994 to April 1995. Part of adolescents from the in-school survey was selected for in-home interviews. Based on the school rosters, a random sample of about 200 students from each high school and feeder school pair was collected to yield the core in-home sample of about 12,000 adolescents. The in-home interviews added special over-samples that included racial/ethnic minorities, physically disabled adolescents, and a genetic sample. These Wave I data produced a total sample size of 20,745 adolescents. Their parents also interviewed in Wave I. Based on the in-school survey, the Add Health network data were constructed.

Thanks to its unique survey design, Add Health provides valuable information suitable for this study. First, the in-home survey contains data on respondents' romantic relationships, such as incidence of romantic relationships in the last 18 months and the number of romantic partners (nominated up to 3) who were either in the sample or not. Second, the in-school survey gathers information on romantic partners' academic achievement, as long as they were in sample or sister schools. Third, the Wave I sample includes a variety of variables that measure respondents' cognitive and noncognitive skills, academic and sexual behaviors, parent-child relationship, school characteristics, family background, and other key demographic characteristics. The analytic sample is restricted to respondents whose information on romantic relationships is available. That is, the analysis excludes the respondents whose romantic partners were not in the sample or sister schools because no information on their academic achievement is obtained.[9] After constructing all measures, the final analytic sample includes 5,199 girls and 5,493 boys.

The NLSY79 is a nationally representative sample of 12,686 young men and women who were 14 to 22 years old when they were first surveyed in 1979. These individuals were interviewed annually through 1994 and are currently interviewed on a biennial basis. Despite its long year coverage, retention rates for the sample are high. Since their first interview, many of the respondents have made transitions from school to work, and from single to married. These data make it possible to study a large sample that represents American men and women born in the late 1950s and 1960s, and living in the United States in 1979. A key feature of this survey is that it gathers information in an event history format, in which dates are collected for the beginning and ending of important life events. Various time-varying variables are detailed in this manner, including union formation, educational attainment, labor market performance, current living arrangement, marriage market characteristics, and some sociodemographic characteristics. The NLSY79 also contains important time-invariant measures such as cognitive and noncognitive skills, family background, and other sociodemographic characteristics.

I construct an event-history file using data from 1980 to 2004 for individuals who were among noninstitutional civilian population. In the analysis, I evaluate patterns of education-based spouse choice for each respondent whose risk of marrying began at age 18. This study only includes respondents at age 18 or younger in 1980 for two reasons. First, inclusion of the older members may create the left-censoring problem. Second, as discussed in the measures section below, measures of cognitive and noncognitive skills are constructed as pre-schooling and pre-labor market factors to reduce reverse causality. The risk of marriage

[9] Alongside limited information on length of the relationships and the number of romantic partner nomination, this restriction may create bias. There is the possibility that the respondents not identified due to these restrictions differ systematically from the sample respondents. Therefore, while Add Health helps increase our understanding of patterns of educational assortative romantic relationship among adolescents, I caution against the generalizability of the results reported here.

is initiated in 1980 to utilize lagged explanatory variables, which were measured from the year prior to the year of interest for the outcome variable. Where information on these explanatory variables is not available, I backfill the missing covariates with retrospective data from the next available year or impute them from the previous available year. After constructing all measures, the final analytic sample size includes 2,145 women and 2,403 men. On average, they contribute 8.01 and 8.92 person-years, respectively, until they marry or drop out of the survey. Therefore, 17,189 person-year observations for women and 21,425 for men are used for analysis.

Measures

Add Health

EDUCATIONAL ASSORTATIVE ROMANTIC RELATIONSHIPS The dependent variable for this analysis is adolescent partner selection by academic achievement. It is categorized as college-bound partner, non-college-bound partner, and single. I treat romantic partners as college-bound if their grade point average (GPA) is 3.0 or higher and as non-college-bound if not (Raley, Crissey, and Muller 2007).[10] If more than two romantic partners were reported, I use the average GPA of these multiple partners.

COGNITIVE AND NONCOGNITIVE SKILLS The key explanatory variables in this study are levels of cognitive and noncognitive skills individuals possess during adolescence. As a measure of cognitive skills, the Wave I in-home survey contains data on the Add Health Picture Vocabulary Test (AHPVT), an abbreviated version of the Peabody Picture Vocabulary Test. Noncognitive skills are measured by several indexes consisting of attitudinal and behavioral aspects of such skills. The Wave I in-home survey provides a subset of the Rotter's locus of control scale and the Rosenberg's self-esteem scale. Locus of

[10] In a sensitivity analysis (not shown), I applied a stricter cutoff point, 3.3, to construct the dependent variable and found similar results to those reported here.

control measures the degree of control individuals feel ranging from external to internal. According to Rotter (1966), individuals who believe that outcomes are due to luck have an external locus of control while individuals who believe that outcomes are due to their own efforts have an internal locus of control. The self-esteem scale measures perceptions of self worth (Rosenberg 1965). I construct a composite index of both scales, which has been commonly employed in past research of the effects of noncognitive skills on socioeconomic outcomes (Carneiro and Heckman 2003).

This study constructs pre-schooling measure of cognitive and noncognitive skills to reduce endogeneity by which unexpected academic achievement might develop both skills valued in schools. Specifically, these measures are constructed by calculating residual values from a regression model where AHPVT or the composite index of noncognitive skills is regressed on age dummies and grade. Then these residual values are standardized to serve as the measures of both skills. We need to be cautious in using these measures because there may be an issue of measurement error. It is unclear how perfectly both measures represent individuals' cognitive and noncognitive skills, but these measures resonate well with the theoretical views described earlier, and as already noted, they have been used frequently in studies of cognitive and noncognitive skills. In addition to constructing these skills as the pre-schooling factor, I introduce both skill measures simultaneously in all models in order to cancel out some of the measurement error problem, given that motivated individuals are more likely to obtain higher test scores.

STUDENT'S OTHER CHARACTERISTICS Since adolescents' academic performance and sexual behavior likely mediate the association between skill differences and adolescent romantic partner choice by academic achievement, I include GPA, school attachment, and incidence of sexual intercourse in the analysis. School attachment is constructed as the sum of three 5-point Likert items, ranging from 3 (low) to 15 (high). These items consist of how close students feel to people at their school, how they feel like they are a part of the school, and how happy they are to be at the school.

PARENT-CHILD RELATIONSHIP Adolescents' skill levels and their romantic behaviors are likely to be governed by parent-child relationship. This relationship can not only influence adolescent skill formation but also constrain incidence of adolescent romantic relationship. I construct two variables that measure parent-child relationship with parental monitoring and parent-child closeness (Harris 1999). Parental monitoring indicates how involved parents are in children's activities, measured by the total count of their activities monitored by parents, ranging from 0 (low) to 7 (high), including curfews, friendships, TV watching, and food and dress choices. Parent-child closeness is measured by adolescents' response to the degree of closeness, satisfaction, warmth, and satisfaction with communication in the parent-child relationship, ranging from 0 (low) to 4 (high). If both parents were present, the average value of the mother and father scores is assigned. If not, a parent-specific value is assigned. I retain missing cases by assigning them the sample mean value and including a missing indicator in the analysis.

SCHOOL CHARACTERISTICS School environment and climate could function as an important constraint for students to form romantic relationships. Following Coleman (1990), this analysis focuses on the collective socialization that is measured by the percentage of white in school, school type (private/public), school location (urban/suburban/rural), school region (Northeast/West/Midwest/South), school-level degree of participation in extracurricular activities, and school mean GPA.

SOCIODEMOGRAPHIC CONTROL VARIABLES I include an array of these control variables in the analysis, including grade (7 to 12), race/ethnicity (white/black/Hispanic/Asian), immigrant generation ($1^{st}/2^{nd}/3^{rd}+$), and family background. Family background covers family structure, parental education, and number of siblings. Family structure is categorized as two-biological parent families, two-parent step families, single-mother families, single-father families, and other families (e.g., foster families). Parental education is measured with the highest level of education either of the parents obtained. It is categorized as less than high school, high school diploma, some college experience, and 4-year college diploma or more, with an indicator of missing observations.

NLSY79

EDUCATION-BASED MARITAL SORTING The dependent variable for this analysis is adult spouse selection by educational attainment. It is an indicator variable of whether a respondent entered a first marriage with 4-year college-graduate, with non-college-graduate, or remained single within a calendar year period.[11]

COGNITIVE AND NONCOGNITIVE SKILLS For the measure of cognitive skills, the NLSY79 provides an aptitude indicator, the full Armed Services Vocational Aptitude Battery (ASVAB) consisting of a series of tests measuring knowledge and skill in areas such as mathematics and language. It was administered to 94 percent of the sample respondents in 1980. A composite score derived from the ASVAB is used to construct an Armed Forces Qualifications Test (AFQT) score, which has been extensively used to measure cognitive skills (Cawley et al. 2000). For the measure of noncognitive skills, a composite index of the locus of control scale and the self-esteem scale is used as that in Add Health. Both scales were administered in 1979 and 1980, respectively. In the almost same fashion as in Add Health, I construct the measures of cognitive and noncognitive skills as pre-schooling and pre-labor market factors to minimize potential biases resulting from the effect of academic and labor market success on both skills. First, I restrict the analytic sample to respondents who were adolescents in 1980 (15 to 18 years old). Second, I calculate residual values from a regression model where the composite index of cognitive or noncognitive skills is regressed

[11] This variable treats cohabitants as single. Whether cohabiting couples can be sorted by educational similarity is of interest in its own right, it is beyond the scope of this study.

on age dummies and years of schooling in 1980. As in Add Health, these residual values are standardized to be used as the measures of both skills.[12]

ECONOMIC POTENTIAL Since men's and women's economic potential has been considered a key factor of entering into a first marriage in general and patterns of educational marital sorting in particular, this analysis includes time-varying measures of educational attainment, earnings, employment status, and homeownership. Educational attainment is measured by highest level of education obtained and categorized as less than high school, high school diploma, some college experience, and 4-year college diploma or more. This study treats GED holders as high school dropouts, following Cameron and Heckman (1993). Annual earnings are measured by summing all wages and business income. To make them a more reliable measure of earnings, they are averaged over 3 years. Nominal earnings data are inflated to 2000 price levels by the implicit deflator of personal consumption expenditures for gross national product. Log-transformed earnings are used in the analysis. Employment status is categorized as not employed, working part-time (less than 1500 hours per year), and working full-time (more than or equal to 1500 hours per year) in a calendar year.

CURRENT LIVING ARRANGEMENT This analysis includes four time-varying measures of current living arrangement: cohabiting; living with child; living with parent; and school enrollment.

MARRIAGE MARKET CHARACTERISTICS The NLSY79 contains several time-varying measures that can be utilized as marriage market characteristics. I construct them with living in an urban area, living in the South, and local unemployment rate (more than or equal to 6 percent or not). Note that they are crude measures by any standard, so one should be cautious

[12] The measures of cognitive skills differ between Add Health and the NLSY79. While the AHPVT score measures verbal ability, the AFQT score also measures mathematical ability. In addition, the NLSY79 contains more items of the locus of control and self-esteem scales. Therefore, the analysis using Add Health may underestimate the role of cognitive and noncognitive skills in educational assortative romantic relationship among adolescents.

in interpreting estimates of these marriage market characteristics (Lichter et al. 1992; Lloyd and South 1996; Lewis and Oppenheimer 2000).

SOCIODEMOGRAPHIC CONTROL VARIABLES This analysis controls for a variety of time-invariant and time-varying measures of sociodemographic variables, such as age, age squared, race/ethnicity (white/black/Hispanic/other), living with two-biological parent families at age 14, parental education at age 14 (less than high school/high school diploma/some college experience/4-year college diploma or more), parental occupation in 1979 (professional or management occupation or not), number of siblings in 1979, living in an urban area at age 14, and living in the South at age 14. A missing indicator of parent occupation is also included in the analysis.

Methods

For the first set of analysis, I employ multinomial logistic regression models using the Add Health data to examine the association between students' cognitive and noncognitive skills and their patterns of educational assortative romantic relationship.[13] Multinomial logistic regression models allows for estimating adolescent romantic partner choice by academic achievement as a function of different underlying mechanisms.[14] The models are expressed as the following equation:

---

[13] Due to Add Health's complex sampling design, the following analysis uses sampling weights, stratification, and clustering to obtain unbiased estimates representing the entire population (Chantala 2001).

[14] This study assumes that patterns of educational assortative mating for both adolescents and adults are distinct with one another. This assumption of irrelevance of independent alternatives (IIA) may be a concern in multinomial logistic models. For each analysis using Add Health or the NLSY79, I ran two standards tests of the IIA assumption and found that the Small-Hsiao test could not reject the null that each category of the dependent variable is independent of other alternatives, while the Hausman-McFadden test was not consistent. Cheng and Long (2007), however, show that it is quite common to find the resulting Hausman-McFadden test negative. In fact, the Small-Hsiao test also has its own caveats, one of which is severe rejection rates when the null is true even in large samples if there are sparse cells in the table of the dependent variable with a binary explanatory variable. Accordingly, Cheng and Long conclude that both tests are not useful for assessing IIA and recommend making the

$$\log\left(\frac{P_{ij}}{1-P_{ij}}\right) = \sum_{m=1}^{M} \beta_m X_{mij}, \tag{1}$$

where P_{ij} is the conditional probability of each pattern of educational assortative romantic relationship (j = 1 for college-bound partner, j = 2 for non-college-bound partner, and j = 3 for single) for a student i. There are m explanatory variables, X, with their coefficients, β_m. This analysis estimates two models separately for boys and girls: the baseline model (Model 1) includes skill differences and sociodemographic controls; and the full model (Model 2) introduces student's other characteristics, parent-child relationship, and school characteristics to address how these constraining and mediating factors influence parameter estimates of skill differences. Note that given the cross-sectional nature of the analytic sample, it is difficult to establish causality. That is, while this analysis examines the influence of skill differences as a pre-schooling factor on adolescent educational assortative partner choice, it is still possible that the causal direction goes the opposite way. Therefore, I focus on the association between the two rather than on building causal models.

For the second set of analysis, I estimate discrete-time event-history models with competing risks using the NLSY79 data to analyze the effect of cognitive and noncognitive skills on patterns of spouse choice by education among adults. These patterns are events to be estimated within a discrete time point, which is each calendar year. When the data are organized as a person-year file, multinomial logistic regression models are equivalent to competing risks models (Allison 1995). In the person-year file, respondents transition to each pattern of education-based marital sorting as competing risks. The functional form of the multinomial discrete-time event-history models is given by:

$$\log\left(\frac{P_{ijt}}{1-P_{ijt}}\right) = \alpha_{tj} + \sum_{m=1}^{M} \beta_m X_{mij} + \sum_{n=1}^{N} \beta_n X_{nij(t-1)}, \tag{2}$$

---

alternatives dissimilar with a careful consideration. I argue that each pattern of education-based mate choice is reasonably distinct, given the prevalence of social segregation in schools by academic achievement among adolescents and educational homogamy among adults.

where P_{ijt} is the conditional probability of transitioning into each pattern of education-based spouse selection (j = 1 for college-graduate spouse, j = 2 for non-college-graduate spouse, and j = 3 for single) for an individual i at year t, given that the individual has remained single or not been censored prior to year t. α_{tj} is a measure of time dependence, estimated by the age and age squared variables to test a nonlinear effect of age. There are m time-invariant explanatory variables with their coefficients, β_m, and n time-varying lagged explanatory variables with their coefficients, β_n. Two models are estimated separately for men and women in this analysis: the baseline model (Model 1) includes skill differences and sociodemographic controls; and the full model (Model 2) adds individuals' economic potential, current living arrangement, and marriage market characteristics to assess the constraining and mediating role of these factors in evaluating the effect of skill differences. For both analyses (one with Add Health and the other with the NLSY79), all coefficients are exponentiated to convert into odds for easier interpretation.

RESULTS

Educational Assortative Romantic Relationships among Adolescents: Add Health

Table 3.1 shows descriptive results for the bivariate relationships between the explanatory variables and patterns of education-based romantic partner choice among adolescents. The last column for each gender reports statistical significance levels of the difference between each pattern by each covariate. First, for girls, dating college-bound partners is more likely among those who have higher levels of cognitive and noncognitive skills, GPA, school attachment, have ever had sexual intercourse, are in a school that are predominantly white, have a higher level of extracurricular activities, or in the West, are in higher grades, white, the 3rd+ generations, come from two-parent families, and have at least one college graduate parent. Because students' characteristics other than skill levels are affected by both skills, they could mediate the bivariate correlation between skill differences and partner selection by academic achievement. Second, dating non-college-bound partners is prevalent among girls who have ever had sex, are in a school that is predominantly white or public, in higher grades, black, the 3rd+ generations, come from step or single-father families, and have parents with high school education. Third, having no romantic relationships is more likely for girls who have a higher level of noncognitive skills (compared to those who date non-college-bound partners), stronger parent-child relationship, are in a school that is in an urban area or in the West, non-white, the 1st or 2nd immigrant generations, come from single-parent families, have parents with less than high school education, and have more siblings. Not surprisingly, higher levels of parental monitoring and parent-child closeness appear to keep girls from having romantic relationships, suggesting its constraining role.

<< Table 3.1 about here >>

Romantic partner sorting on academic achievement for boys shows several differences to that for girls, while most of the bivariate relationships are similar. First, there is no statistical difference in the level of noncognitive skills by educational assortative partner choice. Second, parent-child relationship is strongest among boys who remain single. Third, dating college-bound partners is more likely among boys who are in a school that is in a rural area or in the Northeast. Fourth, family structure appears to have little influence on boys' romantic partner choice, compared to that by girls, except for single-mother and two-biological parent families.

In Table 3.2, I present multinomial logistic regression results for patterns of educational assortative romantic relationship for girls. The baseline model (Model 1) shows that cognitive skills have no association with patterns of romantic partner choice by academic achievement, whereas noncognitive skills are significantly associated with some of these patterns. One standard deviation increase in noncognitive skills is associated with a 15 percent decease in the odds of dating non-college-bound partners over remaining single and with a 12 percent increase in the odds of dating college-bound partners over non-college-bound partners. In the full model (Model 2), however, neither cognitive nor noncognitive skills are significantly associated with girls' romantic partner choice by academic achievement. In both models, the effect of the sociodemographic control variables is generally consistent with the descriptive results. As shown in Model 2, girls in higher grades are more likely to have romantic relationships. Blacks are less likely to date college-bound partners, while Hispanics are more likely to remain single. Having college-educated parent increases the odds of dating college-bound partners, but its effect becomes insignificant in the full model.

<< Table 3.2 about here >>

Model 2 shows that girls' other characteristics, parent-child relationship, and school characteristics moderate and/or constrain the role of noncognitive skills. Girls' own GPA is positively associated with the odds of dating college-bound partners, while school attachment and the experience of sexual intercourse increase the odds of having romantic relationships in general. Parent-child relationship constrains the role of noncognitive skills by its negative effect on the odds of dating college-bound partners (parental monitoring)[15] or on the odds of dating non-college-bound over college-bound partners (parent-child closeness). Going to a predominantly white school is positively associated with dating non-college-bound partners, whereas school-level mean GPA is negatively associated with dating non-college-bound partners. To further explore how the role of noncognitive skills is influenced by these factors, the lower panel in Table 3.2 displays changes in skill coefficients as each factor is introduced in Model 1. Only parent-child relationship accounts for the negative association between noncognitive skills and the odds of dating non-college-bound partners over remaining single. On the other hand, girls' other characteristics, parent-child relationship, and school

[15] This finding seems at odds with the conjecture that parent prefers the more academically achieved as their daughters' dating partners. My interpretation is that as suggested in the descriptive results, the effect of parental monitoring is more salient to girls with non-college-bound partners, who are less likely to be attached to their parent.

characteristics all make their own contribution to explaining the positive association between noncognitive skills and the odds of dating college-bound over non-college-bound partners.

These findings suggest that noncognitive skills play a sorting role in girls' education-based partner selection by decreasing the likelihood of dating non-college-bound partners. Girls' noncognitive skills are not only more valued by their college-bound partners than are cognitive skills, but also increase the likelihood of remaining single if dating non-college-bound partners is given as the only option for having romantic relationships. However, the role of noncognitive skills plays out in an indirect manner to be mediated and constrained by the factors discussed above. Parent-child relationship is particularly salient in this regard because it generally discourages girls to have romantic relationships.

For boys, educational assortative romantic relationship by skill differences displays different patterns compared to that for girls (Table 3.3). In both baseline (Model 1) and full (Model 2) models, cognitive skills are positively associated only with the odds of dating college-bound partners over remaining single, whereas noncognitive skills have no association with romantic partner choice by academic achievement. The effect of the sociodemographic control variables is similar to that shown in the descriptive results. Model 2 shows that boys in higher grades are more likely to have romantic relationships. Blacks are more likely to remain single, while Hispanics are more likely to date non-college-bound partners. Boys from single-mother families are also more likely to date non-college-bound partners. Having parent with some college education increases the likelihood of having romantic relationships.

<< Table 3.3 about here >>

The results from Model 2 suggest that boys' other characteristics do not affect the role of cognitive skills, although they are positively associated with the odds of dating college-bound partners over remaining single. Parent-child relationship and school characteristics—except for urban schools—also have little impact. Meanwhile, as reported in the lower panel in Table 3.3, when each of these factors is introduced separately in Model 1, parent-child relationship accounts for the association between cognitive skills and the odds of dating college-bound partners over remaining single, suggesting that it might keep boys from having romantic relationships. However, the role of cognitive skills is not as much constrained by parent-child relationship as is the role of noncognitive skills for girls, as shown in Model 2.

Taken together, the multinomial logistic regression results indicate that although skill differences play a positive role in educational assortative partner choice in adolescence, the overall association between the two is not that strong. I find that there are marked gender differences in this regard. For girls, a higher level of noncognitive skills increases the likelihood of dating college-bound partners, while it is cognitive skills that increase this likelihood for boys. This perhaps reflects gender differences in a normative attitude toward partner choice that regards "smartness" as more attractive to females than to males. Compared to dating non-college-bound partners, remaining single is a preferred option for girls' with high noncognitive skills, but not for boys regardless of types of skills. In addition, the association between skill differences and partner selection by academic achievement is completely explained by the mediating and constraining factors for girls, but not for boys. This suggests that girls are more likely than boys to be attached to their families and schools and these structured environments outweigh skill differences with respect to girls' education-based partner choice.

Education-Based Marital Sorting among Adults: NLSY79

Table 3.4 reports the means and proportions for each covariate by patterns of education-based spouse selection among adult men and women. First, among women, marrying college graduate is more likely for those who have higher levels of cognitive and noncognitive skills, some college education or more, higher earnings, work full-time, own home, are enrolled in school, live in an urban area, are older, white, lived with two-biological parent families at age 14, had at least one parent with some college education or more at age 14, at least one parent in professional or managerial occupations in 1979, and lived in an urban area at age 14. Second, marrying non-college graduate is more likely for women who have less than high school or high school education, work part-time, cohabit, live in the South or in an area with high unemployment rate, are "other" race/ethnicity, had parent with less than high school or high school education at age 14, had more siblings in 1979. Third, singlehood is more prevalent among women who have some college education, are not employed or work part-time, live with child or parent, are enrolled in school, live in an urban area, are older, black, had parents with less than high school education at age 14, had more siblings in 1979, and lived in an urban area at age 14.

<< Table 3.4 about here >>

For men, the bivariate relationships between each covariate and patterns of marital sorting by education are similar to those for women, with some exceptions. First, there is no statistical difference in having some college education by education-based spouse choice. Second, living with a child is more likely for those who marry non-college graduates. Third, Hispanics are more likely to marry non-college graduates or remain single than to marry college-graduates. Fourth, there is no statistical difference in having had a parent with some college education at age 14 by spouse sorting on educational attainment.

Table 3.5 reports results from the multinomial discrete-time event history analysis for women. The baseline model (Model 1) shows that skill differences are highly predictive of patterns of education-based spouse choice. One standard deviation increase in cognitive skills is significantly associated with a 56 percent increase in the odds of transitioning into marrying a college graduate over remaining single, a 10 percent decrease in the odds of transitioning into marrying a non-college graduate over remaining single, and a 73 percent increase in the odds of transitioning into marrying a college-graduate over a non-college graduate. Also, one standard deviation increase in noncognitive skills is significantly associated with a 30 percent increase in the odds of transitioning into marrying a college graduate over remaining single or marrying a non-college graduate. However, as presented in the full model (Model 2), the effect of cognitive skills becomes insignificant, whereas that of noncognitive skills remains significant. The effect of the sociodemographic control variables generally follows that suggested in the descriptive results. Model 2 shows that as prior research has suggested, there is the nonlinear effect of age whereby the odds of transitioning into marrying a college-graduate increases with age but at a slower rate at later ages. Blacks are more likely to remain single, having had a parent with some college education increases the likelihood of transitioning into marrying a college graduate, having had a parent with college education or lived in an urban area at age 14 decreases the likelihood of transitioning into marrying a non-college graduate.

<< Table 3.5 about here >>

The results from Model 2 suggest that women's cognitive skills have an indirect effect on patterns of education-based spouse sorting, either mediated or constrained by their economic potential, current living arrangement, and the characteristics of marriage market where they live, while their noncognitive skills still have a direct effect even after controlling for these factors. I find that women's educational attainment has a highly significant effect on the odds of transitioning into marrying a college graduate, with the strongest effect for the 4-year college-educated followed by women who have some college education.[16] Homeownership has a positive effect on the odds of entering into first marriage. Living with parent increases the odds of remaining single, while being enrolled in school decreases the odds of transitioning into marrying a non-college graduate. Marriage market characteristics have no

[16] The results show that women's earnings do not have the comparable effect to that of educational attainment. This does not mean the earnings effect is not important for patterns of education-based spouse selection among women. A supplemental analysis (not shown) suggests that the earnings effect is mostly through homeownership.

effect on patterns of education-based spouse selection, but this finding is inconclusive because the measures are not sufficiently refined to detect their role. The lower panel of Table 3.5 shows that when women's economic potential, current living arrangement, and marriage market characteristics are added separately to Model 1, only economic potential and current living arrangement partially account for the effect of cognitive skills. Any of these factors do not explain away the effect of noncognitive skills.

These findings clearly indicate the significant role of skill differences in patterns of spouse selection by educational attainment: women who have higher levels of cognitive and noncognitive skills are more likely to marry college graduates, with each skill playing a distinctive role. The effect of noncognitive skills is particularly strong, suggesting that besides women's economic independence, those with such cultural capital are highly valued by potential college-graduate spouses. In the meantime, cognitive skills increase the likelihood of marrying college graduates in an indirect way to strengthen economic potential, constrained by current living arrangement though.

In Table 3.6, I find that there is a stark gender difference in the association between skill differences and patterns of education-based spouse selection. For men, the baseline model (Model 1) shows the significant effect on these patterns of cognitive skills, but not of noncognitive skills. This finding holds true even in the full model (Model 2) with the exception that the negative effect of cognitive skills on the odds of transitioning into marrying a non-college graduate over remaining single becomes insignificant. Cognitive skills still have a positive impact on the odds of transitioning into marrying a college graduate. The effect of the sociodemographic variables changes as men's economic potential, current living arrangement, and the characteristics of their marriage market are introduced. While the nonlinear effect of age remains significant, most of the other variables lose their explanatory power. For example, having had a parent in professional or managerial occupations in 1979 increases the likelihood of transitioning into marrying college graduate in Model 1, but this positive effect is insignificant in Model 2. This suggests that for men, the effect of family socioeconomic status is not only mediated by their cognitive skills but also by their economic potential.

<< Table 3.6 about here >>

Model 2 reports that men's other characteristics than cognitive skills also have an effect on patterns of spouse sorting by education. College-educated men are more likely to transition into marrying a college graduate or remain single if marrying a non-college graduate is the only option for them. Earnings and home ownership increase the likelihood of entering into first marriage regardless of spouse's educational attainment. Cohabiting and being enrolled in school decrease the likelihood of transitioning into marrying a non-college graduate, whereas living with child increases this likelihood. Living with parent increases the likelihood of remaining single. Marriage market characteristics have little effect on education-based spouse choice, except for living in an urban area. However, as indicated in the lower panel of Table 3.6, when these factors are introduced separately in Model 1, none of them affect the positive effect of cognitive skills on marrying a college graduate, although men's economic potential and current living arrangement reduce their effect size.

In summary, the results from the multinomial discrete-time event-history models demonstrate that skill differences play a significant role in education-based marital sorting. While economic potential and current living arrangement also have their own impact on this sorting process, their mediating role is limited. The finding that skill differences have the independent effect illuminates the mechanism by which familial resources are associated with patterns of spouse selection by educational attainment, given that one's skill levels are highly dependent on family socioeconomic status and parenting. In addition, the results reveal clear gender differences in the role of cognitive and noncognitive skills. For women, marrying college graduates depends significantly on both skills, although the effect of cognitive skills is indirect as their economic potential mediates it. For men, marrying college graduates depends more on cognitive skills than on noncognitive skills. This finding suggests that socially valued noncognitive skills as cultural capital primarily shape the preference for educational similarity between spouses for women, while it is cognitive skills that do so for men. This divergence also hints at the normative dimension of mate selection processes by which well-educated women may be more attracted to men with high cognitive skills and well-educated men may be more attracted to women with high noncognitive skills.

DISCUSSION AND CONCLUSION

Students of educational assortative mating have long pointed out its relations to socioeconomic inequality (Mare 1991; Schwartz and Mare 2005). They conjecture that the growing educational resemblance of spouses indicates an important facet of social closure, because it reflects a decrease in interactions between social status groups, which is likely to lead to differentials in socioeconomic outcomes in the next generation. Identifying the linkage between the current prevalence of educational homogamy and intergenerational mobility, however, would not be immediately available without data on the next generations. In this paper, I address the association between educational assortative mating and socioeconomic inequality by taking the alternative approach that is based on two observations from the literature on martial sorting. First, it is difficult to differentiate whether one's own education reflects his or her family background or the meritocratic characteristic of the education system. Second, there has been a lack of research interest in the intimate nature of mate selection processes. I argue that the role individual skill differences play in patterns of educational assortative mating helps to clarify its connection to socioeconomic inequality, given the sheer influence of familial resources on developing cognitive and noncognitive skills.

The preceding analyses demonstrate that skill differences are predictive of education-based mate selection. Skill differences are positively associated with the likelihood of transitioning into marrying college graduates in adulthood. For women, higher levels of cognitive and noncognitive skills increase the yearly probability that their spouse is college-educated; cognitive skills have a more indirect effect as it is mediated by economic potential, while noncognitive skills have a more direct effect. For men, this probability depends mostly on a level of cognitive skills. The corresponding association in adolescence is not as strong as in adulthood but is still suggestive. For girls, a higher level of noncognitive skills is indirectly associated with the probability of dating college-bound partners or remaining single, implying the strong moderating and/or mediating role of own academic achievement, parent-child relationship, school characteristics. For boys, a higher level of cognitive skills directly promotes the probability of dating college-bound partners, under the condition that boys are less constrained than girls from having romantic relationships.

These results clearly indicate that there is a gender difference in what type of skills is more important in predicting the probability of transitioning into marrying the college-educated or dating college-bound partners. It is primarily noncognitive skills for women, while it is cognitive skills for men, which are responsible for their preference for educational similarity between spouses or romantic partners. This pattern also reflects that in addition to the strong educational resemblance of spouses or romantic partners, noncognitive skills for women and cognitive skills for men are more valued by potential well-educated mates in the intimate process of mate selection.

Despite this gender difference, this study has straightforward implications for the linkage between educational assortative mating and socioeconomic inequality. While lending support to the prior research that has shown that marital sorting by education indicates marriage as a marker of prestige (Cherlin 2004), this analysis further specifies a way in which family socioeconomic status affects education-based spouse choice. The findings suggest that the intergenerational transmission of familial resources takes a more complex form, as it relates to not simply investing in children's educational attainment but also strengthening their cognitive and noncognitive skills. Although to lesser degree, this pattern is also associated with educational assortative romantic relationships in adolescence. Taken together, the role of skill differences in education-based mate selection illuminates that socioeconomic inequality operates even at an intimate level of the mating process.

While the present study identifies cognitive and noncognitive skills as an important link between family background and individual patterns of educational assortative mating, there are other critical questions that warrant attention. First, even if a large body of literature has documented skill differences by family socioeconomic status, the mechanism by which family background affects skill formation needs to be uncovered by gaining more knowledge about parenting behaviors and the gene-environment interactions. In this vein, developing more reliable measures of cognitive and noncognitive skills can benefit from utilizing multiple indices of aptitude and behavioral traits (e.g., child problem behaviors). Although this study constructs both skills as pre-schooling and/or pre-labor market measures, they still cannot rule out the potential of measurement error. Second, given that educational assortative mating involves a matching process between two sexes, explanations for the gender difference in the role of skill differences require to distinguish between the "readiness" of individuals with higher levels of skills for mating the well-educated and their "attractiveness" to potential well-educated mate candidates. This study provides an explanation for this gender difference with the assumption that the process of education-based mate selection is likely to go both directions; however, it is inconclusive until more research that takes the characteristics of both sexes into account is done. Lastly, but not least, examinations of how skill differences are associated with sorting by education among cohabiting couples or racial groups should be given a priority. Although data limitations prevent this study from conducting separate analyses by mating types or race/ethnicity, such analyses are needed to understand a more comprehensive picture of the role of skill differences in individual patterns of educational assortative mating.

Table 3.1. Means and Proportions of Analysis Variables by Educational Assortative Romantic Relationships among Adolescents, Add Health

	Girls				Boys			
Variable	College-Bound Partner (1)	Non-College-Bound Partner (2)	Single (3)	Significant Difference (*p* <.05)	College-Bound Partner (1)	Non-College-Bound Partner (2)	Single (3)	Significant Difference (*p* <.05)
Dependet variable	.14	.17	.69		.21	.14	.66	
Student's Characteristics								
Cognitive skills (Add Health PVT score)	.31	.14	.01	1 > 2 > 3	.37	.21	.15	1 > 2,3
Noncognitive skills (Rotter/Rosenberg scale)	.03	-.15	-.02	1,3 > 2	.23	.11	.16	ns
GPA	3.19	2.85	2.99	1 > 3 > 2	2.87	2.50	2.76	1 > 3 > 2
School attachment	12.01	11.48	11.55	1 > 2,3	11.82	11.23	11.42	1 > 2,3
Ever had sex	.36	.43	.14	1,2 > 3	.43	.54	.20	2 > 1 > 3
Parent-Child Relationship								
Parent's monitoring	1.59	1.85	2.20	3 > 2 > 1	1.71	1.64	2.11	3 > 1,2
Parent-child closeness	3.20	3.07	3.31	3 > 1 > 2	3.30	3.22	3.37	3 > 1,2
School Characteristics								
Percent white	71.57	69.52	61.83	1,2 > 3	71.21	64.19	64.11	1 > 2,3
Percent extracurricular activities	85.74	84.11	84.57	1 > 2,3	85.13	83.71	84.42	1 > 2
Public	.91	.96	.93	2 > 1,3	.93	.96	.92	ns
Urban	.21	.16	.28	3 > 1,2	.16	.20	.26	3 > 1
Rural	.22	.21	.15	ns	.23	.20	.15	1 > 3
Suburban (reference)	.58	.63	.57	ns	.62	.60	.59	ns
West	.17	.11	.19	1,3 > 2	.14	.13	.17	ns
Midwest	.30	.36	.30	ns	.27	.33	.30	ns
South	.36	.39	.37	ns	.43	.44	.38	ns
Northeast (reference)	.16	.13	.14	ns	.16	.11	.15	1 > 2
School mean GPA	2.90	2.76	2.84	1 > 3 > 2	2.88	2.75	2.85	1,3 > 2
Sociodemographic Controls								
Grade	9.52	9.50	8.81	1,2 > 3	9.59	9.73	9.00	1,2 > 3
Black	.05	.14	.16	2,3 > 1	.09	.16	.13	2,3 > 1
Hispanic	.05	.06	.14	3 > 1,2	.07	.13	.12	2,3 > 1
Asian	.05	.03	.06	3 > 2	.03	.03	.06	3 > 1,2
White (reference)	.85	.77	.64	1 > 2 > 3	.80	.68	.68	1 > 2,3
1st generation	.02	.03	.08	3 > 1,2	.03	.02	.06	3 > 1,2
2nd generation	.08	.09	.12	3 > 1	.08	.11	.13	3 > 1
3rd+ generation (reference)	.90	.88	.81	1,2 > 3	.90	.86	.81	1,2 > 3
Step	.19	.20	.14	1,2 > 3	.18	.20	.17	ns
Mother only	.14	.19	.21	3 > 1	.13	.25	.17	2,3 > 1
Father only	.01	.03	.02	2,3 > 1	.03	.03	.04	ns
Other family structure	.03	.03	.03	ns	.03	.05	.03	ns
Two biological parents (reference)	.63	.55	.60	1 > 2	.63	.48	.60	1,3 > 2
Less than high school	.08	.07	.13	3 > 1,2	.07	.12	.11	2,3 > 1
High school graduate (reference)	.25	.36	.31	2 > 3 > 1	.29	.31	.33	ns
Some college	.23	.24	.21	ns	.24	.22	.18	1 > 3
College graduate	.43	.30	.32	1 > 2,3	.38	.32	.34	1 > 2
Number of siblings	1.33	1.33	1.54	3 > 1,2	1.37	1.34	1.48	3 > 1
N		5,199				5,493		

Notes: Nonsignificant differences between each category of the dependent variable are indicated by ns; Means and proportions are weighted and adjusted for design effects.

Table 3.2. Multinomial Logistic Regression of Educational Assortative Romantic Relationships for Girls, Add Health

	Model 1			Model 2		
	College-Bound vs. Single	Non-College-Bound vs. Single	College-Bound vs. Non-College-Bound	College-Bound vs. Single	Non-College-Bound vs. Single	College-Bound vs. Non-College-Bound
Student's Characteristics						
Cognitive skills	1.108	1.042	1.063	1.038	1.093	0.949
(Add Health PVT score)	(.081)	(.063)	(.084)	(.081)	(.069)	(.078)
Noncognitive skills	.954	.850 **	1.123 *	.875	.924	.947
(Rotter/Rosenberg scale)	(.044)	(.041)	(.061)	(.060)	(.056)	(.071)
GPA				1.446 ***	.940	1.538 ***
				(.132)	(.079)	(.161)
School attachment				1.121 ***	1.083 ***	1.035
				(.036)	(.023)	(.037)
Ever had sex				3.522 ***	3.524 ***	.999
				(.607)	(.506)	(.192)
Parent-Child Relationship						
Parent's monitoring				.877 **	1.001	.876 *
				(.040)	(.036)	(.046)
Parent-child closeness				.844	.737 ***	1.146
				(.080)	(.055)	(.106)
School Characteristics						
Percent white				.998	1.008 *	.989 **
				(.003)	(.004)	(.004)
Percent extracurricular activities				1.033	1.027 *	1.006
				(.017)	(.013)	(.014)
Public				1.402	1.686	.832
				(.392)	(.547)	(.259)
Urban				.932	.757	1.230
				(.163)	(.113)	(.235)
Rural				1.346	1.017	1.323
				(.322)	(.145)	(.321)
Suburban (reference)						
West				.887	.804	1.104
				(.196)	(.173)	(.290)
Midwest				.849	.861	.986
				(.182)	(.154)	(.211)
South				1.051	1.000	1.051
				(.207)	(.172)	(.229)
Northeast (reference)						
School mean GPA				2.137	.167 ***	12.811 ***
				(.844)	(.060)	(4.969)
Sociodemographic controls						
Grade	1.334 ***	1.336 ***	.999	1.243 ***	1.187 ***	1.047
	(.060)	(.056)	(.047)	(.053)	(.056)	(.049)
Black	.292 ***	.715	.408 **	.269 ***	.686	.392 **
	(.078)	(.139)	(.114)	(.084)	(.137)	(.117)
Hispanic	.394 ***	.346 ***	1.138	.447 ***	.478 **	.935
	(.083)	(.092)	(.378)	(.099)	(.121)	(.303)
Asian	.895	.535	1.675	.810	.764	1.061
	(.252)	(.187)	(.595)	(.245)	(.300)	(.421)
White (reference)						

(continued on the next page)

Table 3.2. *(continued)*

	Model 1			Model 2		
	College-Bound vs. Single	Non-College-Bound vs. Single	College-Bound vs. Non-College-Bound	College-Bound vs. Single	Non-College-Bound vs. Single	College-Bound vs. Non-College-Bound
1st generation	.334 **	.570	.586	.411 *	.732	.562
	(.118)	(.176)	(.256)	(.141)	(.238)	(.253)
2nd generation	.759	1.163	.653	.846	1.340	.631
	(.147)	(.274)	(.170)	(.167)	(.338)	(.174)
3rd+ generation (reference)						
Step	1.424 **	1.563 *	.911	1.269	1.172	1.083
	(.189)	(.272)	(.174)	(.202)	(.233)	(.241)
Mother only	.866	1.033	.838	.818	.974	.840
	(.148)	(.139)	(.167)	(.150)	(.135)	(.173)
Father only	.491	1.026	.479	.468	.891	.525
	(.187)	(.324)	(.196)	(.181)	(.329)	(.241)
Other family structure	1.193	1.328	.898	1.363	1.369	.995
	(.415)	(.402)	(.348)	(.510)	(.440)	(.434)
Two biological parents (reference)						
Less than high school	1.125	.623 **	1.806	1.107	.614 **	1.802
	(.293)	(.105)	(.554)	(.294)	(.112)	(.557)
High school graduate (reference)						
Some college	1.222	.902	1.355	1.200	.972	1.235
	(.197)	(.120)	(.242)	(.200)	(.148)	(.225)
College graduate	1.482 *	.795	1.864 **	1.380	1.009	1.368
	(.242)	(.113)	(.381)	(.244)	(.156)	(.300)
Number of siblings	.916	.915	1.001	.968	.930	1.042
	(.048)	(.051)	(.082)	(.044)	(.049)	(.069)
F-statistics		5.95 ***			7.94 ***	
(df)		(32,90)			(64,58)	
N		5,199			5,199	
Changes in Skill Coefficients						
Model 1 + Student's other characteristics						
Cognitive skills	1.079	1.094	.986			
Noncognitive skills	.830 **	.874 *	.950			
Model 1 + Parent-child relationship						
Cognitive skills	1.077	1.023	1.053			
Noncognitive skills	.998	.935	1.067			
Model 1 + School characteristics						
Cognitive skills	1.091	1.067	1.022			
Noncognitive skills	.938	.856 **	1.096			

Notes: All coefficients are expressed as odds ratios; Robust standard errors in parentheses; All estimates are weighted and adjusted for design effects; Missing indicators for parent-child closeness and parental education are included but not shown.
* p<0.05; ** p<0.01; *** p<0.001 (two-tailed tests).

Table 3.3. Multinomial Logistic Regression of Educational Assortative Romantic Relationships for Boys, Add Health

	Model 1			Model 2		
	College-Bound vs. Single	Non-College-Bound vs. Single	College-Bound vs. Non-College-Bound	College-Bound vs. Single	Non-College-Bound vs. Single	College-Bound vs. Non-College-Bound
Student's Characteristics						
Cognitive skills	1.141 *	1.050	1.088	1.172 *	1.134	1.034
(Add Health PVT score)	(.068)	(.072)	(.082)	(.071)	(.085)	(.091)
Noncognitive skills	1.033	.947	1.091	.942	1.034	.911
(Rotter/Rosenberg scale)	(.063)	(.064)	(.093)	(.070)	(.081)	(.088)
GPA				1.153 *	0.726 ***	1.589 ***
				(.076)	(.057)	(.150)
School attachment				1.100 ***	1.038	1.060
				(.027)	(.027)	(.033)
Ever had sex				3.306 ***	3.520 ***	.939
				(.400)	(.489)	(.137)
Parent-Child Relationship						
Parent's monitoring				.947	.943	1.004
				(.035)	(.034)	(.049)
Parent-child closeness				.838	.802 *	1.044
				(.075)	(.085)	(.110)
School Characteristics						
Percent white				1.000	1.002	.997
				(.003)	(.004)	(.004)
Percent extracurricular activities				1.030	1.042 *	.989
				(.018)	(.019)	(.017)
Public				2.034	1.710	1.189
				(.945)	(.665)	(.457)
Urban				.682 *	.771	.885
				(.127)	(.133)	(.181)
Rural				1.188	0.963	1.233
				(.181)	(.179)	(.205)
Suburban (reference)						
West				1.005	1.353	.743
				(.241)	(.368)	(.163)
Midwest				.856	1.376	.622 *
				(.177)	(.385)	(.122)
South				1.199	1.543	.777
				(.243)	(.380)	(.144)
Northeast (reference)						
School mean GPA				1.794	.193 ***	9.313 ***
				(.610)	(.073)	(3.528)
Sociodemographic controls						
Grade	1.224 ***	1.294 ***	.946	1.117 **	1.144 **	.976
	(.045)	(.052)	(.041)	(.045)	(.048)	(.051)
Black	.726	1.045	.694	.594 *	.669 *	.887
	(.137)	(.236)	(.149)	(.131)	(.135)	(.231)
Hispanic	.807	1.483	.544 *	.979	1.709 *	.573 *
	(.144)	(.356)	(.134)	(.199)	(.376)	(.153)
Asian	.697	.772	.903	.762	.983	.775
	(.156)	(.299)	(.390)	(.209)	(.379)	(.343)
White (reference)						

(continued on the next page)

Table 3.3. *(continued)*

	Model 1			Model 2		
	College-Bound vs. Single	Non-College-Bound vs. Single	College-Bound vs. Non-College-Bound	College-Bound vs. Single	Non-College-Bound vs. Single	College-Bound vs. Non-College-Bound
1st generation	.547 *	.296 ***	1.850	.662	.360 ***	1.837
	(.155)	(.098)	(.842)	(.181)	(.105)	(.752)
2nd generation	.613 ***	.809	.758	.689 **	.942	.732
	(.083)	(.175)	(.186)	(.097)	(.202)	(.176)
3rd+ generation (reference)						
Step	1.069	1.527 **	.700	1.114	1.138	.979
	(.136)	(.235)	(.127)	(.166)	(.219)	(.219)
Mother only	.822	1.804 ***	.456 ***	.842	1.545 **	.545 ***
	(.112)	(.286)	(.076)	(.116)	(.252)	(.094)
Father only	.888	1.113	.798	.925	.901	1.026
	(.232)	(.345)	(.288)	(.257)	(.295)	(.364)
Other family structure	1.231	2.033 *	.605	.963	1.267	.761
	(.362)	(.617)	(.217)	(.296)	(.399)	(.278)
Two biological parents (reference)						
Less than high school	.967	1.151	.840	.877	1.046	.839
	(.204)	(.233)	(.226)	(.198)	(.227)	(.238)
High school graduate (reference)						
Some college	1.473 **	1.316	1.119	1.462 *	1.509 *	.969
	(.206)	(.200)	(.192)	(.214)	(.239)	(.173)
College graduate	1.226	1.160	1.057	1.277	1.614 ***	.791
	(.142)	(.172)	(.173)	(.161)	(.228)	(.127)
Number of siblings	.984	.957	1.028	1.004	.973	1.032
	(.034)	(.051)	(.060)	(.040)	(.055)	(.063)
F-statistics		5.48 ***			11.08 ***	
(df)		(32,90)			(64,58)	
N		5,493			5,493	
Changes in Skill Coefficients						
Model 1 + Student's other characteristics						
Cognitive skills	1.172 **	1.150 *	1.019			
Noncognitive skills	.923	.996	.927			
Model 1 + Parent-child relationship						
Cognitive skills	1.117	1.022	1.093			
Noncognitive skills	1.054	.996	1.059			
Model 1 + School characteristics						
Cognitive skills	1.158 *	1.055	1.097			
Noncognitive skills	1.014	.956	1.060			

Notes: All coefficients are expressed as odds ratios; Robust standard errors in parentheses; All estimates are weighted and adjusted for design effects; Missing indicators for parent-child closeness and parental education are included but not shown.
* $p<0.05$; ** $p<0.01$; *** $p<0.001$ (two-tailed tests).

Table 3.4. Means and Proportions of Analysis Variables by Education-Based Marital Sorting among Adults, NLSY79

Variable	Women				Men			
	College Graduate Spouse (1)	Non-College Graduate Spouse (2)	Single (3)	Significant Difference (*p*<.05)	College Graduate Spouse (1)	Non-College Graduate Spouse (2)	Single (3)	Significant Difference (*p*<.05)
Dependent variable	.02	.06	.92		.01	.05	.94	
Skill Characteristics								
Cognitive skills (AFQT score)	.73	-.06	-.14	1 > 2 > 3	.91	.00	-.02	1 > 2,3
Noncognitive skills (Rotter/Rosenberg scale)	.35	-.12	-.09	1 > 2,3	.37	.05	.06	1 > 2,3
Economic Potential								
Less than high school	.02	.21	.22	2,3 > 1	.04	.30	.32	2,3 > 1
High school graduate (reference)	.12	.48	.39	2 > 3 > 1	.14	.43	.36	2 > 3 > 1
Some college	.31	.23	.26	1,3 > 2	.17	.20	.20	ns
College graduate	.55	.09	.13	1 > 3 > 2	.65	.08	.11	1 > 3 > 2
Log earnings	9.57	8.32	7.62	1 > 2 > 3	10.15	9.43	8.59	1 > 2 > 3
Not employed (reference)	.03	.11	.20	3 > 2 > 1	.02	.04	.11	3 > 1,2
Part-time	.27	.38	.38	2,3 > 1	.17	.24	.36	3 > 1,2
Full-time	.70	.51	.43	1 > 2 > 3	.81	.71	.53	1 > 2 > 3
Own home	.16	.11	.08	1 > 2 > 3	.26	.13	.08	1 > 2 > 3
Current Living Arrangement								
Cohabiting	.13	.17	.11	2 > 3	.11	.14	.11	2 > 3
Child present	.07	.24	.32	3 > 2 > 1	.05	.16	.07	2 > 1,3
Living with parent	.27	.41	.48	3 > 2 > 1	.14	.37	.53	3 > 2 > 1
Enrolled in school	.23	.16	.23	1,3 > 2	.16	.10	.19	3 > 2
Marriage Market Characteristics								
Urban	.87	.77	.84	1,3 > 2	.88	.75	.82	1 > 3 > 2
South	.32	.40	.39	2 > 1	.40	.39	.36	2 > 3
Local unemployment rate (≥ 6%)	.54	.71	.62	2 > 3 > 1	.43	.66	.63	2,3 > 1
Sociodemographic Controls								
Age	25.74	22.83	25.23	1,3 > 2	27.28	24.21	25.52	1 > 3 > 2
Black	.15	.25	.42	3 > 2 > 1	.17	.22	.35	3 > 1,2
Hispanic	.14	.16	.17	ns	.08	.17	.16	2,3 > 1
Other race/ethnicity	.10	.11	.07	2 > 3	.11	.13	.10	2 > 3
White (reference)	.61	.47	.34	1 > 2 > 3	.64	.47	.39	1 > 2 > 3
Two-biological parents at age 14	.79	.68	.60	1 > 2 > 3	.80	.67	.64	1 > 2,3
Parental education at age 14:								
Less than high school	.14	.35	.37	2,3 > 1	.11	.35	.33	2,3 > 1
High school graduate (reference)	.31	.44	.38	2 > 3 > 1	.34	.43	.39	2 > 1
Some college	.18	.10	.11	1 > 2,3	.16	.11	.13	ns
College graduate	.37	.11	.14	1 > 3 > 2	.40	.12	.16	1 > 3 > 2
Parental occupation in 1979: Professional/management	.42	.18	.18	1 > 2,3	.46	.16	.18	1 > 2,3
Number of siblings in 1979	2.95	3.67	3.83	2,3 > 1	2.71	3.73	3.80	2,3 > 1
Urban at age 14	.84	.77	.84	1,3 > 2	.84	.76	.81	1,3 > 2
South at age 14	.31	.39	.37	ns	.34	.37	.35	ns
N (person years)		17,189				21,425		

Notes: Nonsignificant differences between each category of the dependent variable are indicated by ns.

Table 3.5. Multinomial Discrete-Time Event History Analysis of Education-Based Marital Sorting for Women, NLSY79

	Model 1			Model 2		
	College vs. Single	Non-College vs. Single	College vs. Non-College	College vs. Single	Non-College vs. Single	College vs. Non-College
Skill Characteristics						
Cognitive skills	1.557 ***	.898 **	1.734 ***	1.105	.922	1.198
(AFQT score)	(.126)	(.037)	(.155)	(.098)	(.040)	(.117)
Noncognitive skills	1.304 ***	1.007	1.295 ***	1.245 ***	1.001	1.243 **
(Rotter/Rosenberg scale)	(.083)	(.033)	(.091)	(.080)	(.034)	(.089)
Economic Potential						
Less than high school				.611	.981	.623
				(.260)	(.088)	(.270)
High school graduate (reference)						
Some college				2.914 ***	1.018	2.861 ***
				(.640)	(.093)	(.672)
College graduate				5.949 ***	.785	7.578 ***
				(1.429)	(.109)	(2.062)
Log earnings				1.086	1.055 *	1.030
				(.074)	(.022)	(.073)
Not employed (reference)						
Part-time				1.099	1.217	.903
				(.482)	(.176)	(.415)
Full-time				1.454	1.499 *	.970
				(.687)	(.247)	(.482)
Own home				1.518 *	1.963 ***	.773
				(.284)	(.227)	(.166)
Current Living Arrangement						
Cohabiting				1.148	1.186	.968
				(.219)	(.113)	(.203)
Child present				.642	1.078	.595
				(.171)	(.099)	(.167)
Living with parent				.594 ***	.544 ***	1.093
				(.093)	(.042)	(.189)
Enrolled in school				.898	.543 ***	1.653 **
				(.149)	(.054)	(.314)
Marriage Market Characteristics						
Urban				.890	.869	1.024
				(.174)	(.074)	(.215)
South				.878	1.034	.849
				(.194)	(.144)	(.217)
Local unemployment rate				1.075	1.118	.961
(≥ 6%)				(.141)	(.085)	(.144)
Sociodemographic Controls						
Age	3.790 ***	1.151 *	3.292 ***	2.030 ***	.887	2.289 ***
	(.548)	(.067)	(.509)	(.327)	(.057)	(.394)
Age squared	.976 ***	.996 ***	.980 ***	.986 ***	1.000	.986 ***
	(.003)	(.001)	(.003)	(.003)	(.001)	(.003)
Black	.395 ***	.414 ***	.954	.430 ***	.555 ***	.774
	(.082)	(.040)	(.217)	(.094)	(.058)	(.186)
Hispanic	.761	.687 ***	1.107	.848	.809 *	1.048
	(.143)	(.066)	(.231)	(.170)	(.081)	(.232)
Other race/ethnicity	.974	1.065	.915	1.042	1.139	.914
	(.204)	(.114)	(.211)	(.222)	(.125)	(.215)

(continued on the next page)

Table 3.5. *(continued)*

	Model 1			Model 2		
	College vs. Single	Non-College vs. Single	College vs. Non-College	College vs. Single	Non-College vs. Single	College vs. Non-College
White (reference)						
Two-biological parents at age 14	1.377 *	1.124	1.225	1.142	1.247 **	.916
	(.217)	(.081)	(.210)	(.185)	(.092)	(.161)
Parental education at age 14:						
Less than high school	.699	.975	.717	.870	.976	.891
	(.144)	(.076)	(.156)	(.185)	(.077)	(.200)
High school graduate (reference)						
Some college	1.947 ***	.785 *	2.480 ***	1.868 ***	.818	2.282 ***
	(.353)	(.086)	(.515)	(.346)	(.092)	(.484)
College graduate	1.370	.607 ***	2.259 ***	1.093	.714 **	1.530 *
	(.238)	(.073)	(.467)	(.188)	(.088)	(.317)
Parental occupation in 1979:	1.215	1.000	1.215	1.092	1.053	1.038
Professional/management	(.183)	(.095)	(.212)	(.165)	(.102)	(.182)
Number of siblings in 1979	.990	1.005	.985	1.014	1.001	1.012
	(.031)	(.014)	(.033)	(.033)	(.014)	(.035)
Urban at age 14	1.180	.793 **	1.488 *	1.207	.807 **	1.496 *
	(.199)	(.062)	(.273)	(.208)	(.066)	(.280)
South at age 14	1.298	1.285 ***	1.010	1.338	1.255	1.066
	(.178)	(.090)	(.153)	(.294)	(.174)	(.272)
-2 Log likelihood		10,302.87			9,893.89	
N (person years)		17,189			17,189	
Changes in Skill Coefficients						
Model 1 + Economic potential						
Cognitive skills	1.150	.905 *	1.270 *			
Noncognitive skills	1.257 ***	.993	1.266 **			
Model 1 + Current living arrangement						
Cognitive skills	1.440 ***	.934	1.541 ***			
Noncognitive skills	1.282 ***	1.013	1.266 ***			
Model 1 + Marriage market characteristics						
Cognitive skills	1.555 ***	.898 **	1.731 ***			
Noncognitive skills	1.303 ***	1.010	1.290 ***			

Notes: All coefficients are expressed as odds ratios; Standard errors in parentheses; A missing indicator for parental occupation is included but not shown.

* p<0.05; ** p<0.01; *** p<0.001 (two-tailed tests).

Table 3.6. Multinomial Discrete-Time Event History Analysis of Education-Based Marital Sorting for Men, NLSY79

	Model 1			Model 2		
	College vs. Single	Non-College vs. Single	College vs. Non-College	College vs. Single	Non-College vs. Single	College vs. Non-College
Skill Characteristics						
Cognitive skills	1.920 ***	.925 *	2.075 ***	1.209 *	.945	1.280 *
(AFQT score)	(.170)	(.035)	(.198)	(.117)	(.040)	(.134)
Noncognitive skills	1.127	1.050	1.074	1.041	1.015	1.025
(Rotter/Rosenberg scale)	(.076)	(.036)	(.081)	(.074)	(.035)	(.080)
Economic Potential						
Less than high school				.441 *	.888	.496
				(.168)	(.073)	(.193)
High school graduate (reference)						
Some college				1.705 *	1.063	1.604
				(.416)	(.102)	(.416)
College graduate				5.437 ***	.623 **	8.721 ***
				(1.296)	(.090)	(2.392)
Log earnings				1.314 **	1.251 ***	1.050
				(.127)	(.046)	(.107)
Not employed (reference)						
Part-time				.819	.771	1.062
				(.412)	(.139)	(.564)
Full-time				1.001	1.125	.890
				(.515)	(.215)	(.486)
Own home				1.886 ***	1.588 ***	1.188
				(.318)	(.170)	(.232)
Current Living Arrangement						
Cohabiting				.867	.466 ***	1.862 *
				(.204)	(.056)	(.487)
Child present				1.858	4.150 ***	.448 *
				(.646)	(.498)	(.163)
Living with parent				.340 ***	.431 ***	.789
				(.070)	(.034)	(.173)
Enrolled in school				1.065	.574 ***	1.855 **
				(.218)	(.070)	(.436)
Marriage Market Characteristics						
Urban				.896	.742 ***	1.208
				(.192)	(.062)	(.275)
South				1.325	1.157	1.145
				(.275)	(.153)	(.277)
Local unemployment rate				.792	1.030	.769
(≥ 6%)				(.110)	(.077)	(.120)
Sociodemographic Controls						
Age	4.778 ***	1.473 ***	3.244 ***	2.206 ***	.990	2.229 ***
	(.799)	(.087)	(.573)	(.402)	(.064)	(.429)
Age squared	.973 ***	.992 ***	.981 ***	.985 ***	.998	.987 ***
	(.003)	(.001)	(.003)	(.003)	(.001)	(.003)
Black	.801	.404 ***	1.981 **	.794	.483 ***	1.642
	(.177)	(.040)	(.475)	(.185)	(.050)	(.416)
Hispanic	.679	.836	.813	.776	.915	.849
	(.172)	(.081)	(.219)	(.205)	(.094)	(.238)
Other race/ethnicity	.817	.946	.864	.910	.959	.948
	(.174)	(.096)	(.202)	(.199)	(.100)	(.227)

(continued on the next page)

Table 3.6. *(continued)*

	Model 1			Model 2		
	College vs. Single	Non-College vs. Single	College vs. Non-College	College vs. Single	Non-College vs. Single	College vs. Non-College
White (reference)						
Two-biological parents at age 14	1.405 *	.993	1.416	1.236	1.075	1.149
	(.240)	(.071)	(.260)	(.216)	(.079)	(.216)
Parental education at age 14:						
Less than high school	.692	1.009	.685	.833	1.009	.826
	(.167)	(.081)	(.173)	(.207)	(.083)	(.215)
High school graduate (reference)						
Some college	1.054	.780 *	1.350	.809	.780 *	1.037
	(.211)	(.084)	(.304)	(.168)	(.086)	(.240)
College graduate	1.219	.652 ***	1.870 **	.820	.743 *	1.103
	(.217)	(.077)	(.395)	(.149)	(.091)	(.237)
Parental occupation in 1979:	1.416 *	.906	1.562 *	1.200	.936	1.283
Professional/management	(.231)	(.093)	(.297)	(.194)	(.097)	(.243)
Number of siblings in 1979	.959	1.011	.948	1.005	1.007	.998
	(.035)	(.014)	(.037)	(.036)	(.014)	(.038)
Urban at age 14	1.280	.839 *	1.527 *	1.425	.968	1.472
	(.230)	(.065)	(.296)	(.265)	(.079)	(.296)
South at age 14	1.634 ***	1.319 ***	1.239	1.143	1.094	1.045
	(.234)	(.095)	(.197)	(.244)	(.147)	(.259)
-2 Log likelihood		10,396.22			9,618.21	
N (person years)		21,425			21,425	
Changes in Skill Coefficients						
Model 1 + Economic potential						
Cognitive skills	1.264 *	.934	1.354 **			
Noncognitive skills	1.063	1.025	1.037			
Model 1 + Current living arrangement						
Cognitive skills	1.791 ***	.967	1.851 ***			
Noncognitive skills	1.093	1.032	1.059			
Model 1 + Marriage market characteristics						
Cognitive skills	1.895 ***	.927 *	2.044 ***			
Noncognitive skills	1.122	1.054	1.065			

Notes: All coefficients are expressed as odds ratios; Standard errors in parentheses; A missing indicator for parental occupation is included but not shown.

* p<0.05; ** p<0.01; *** p<0.001 (two-tailed tests).

CHAPTER 4

The Early Socioeconomic Effects of Teenage Childbearing:
A Propensity Score Matching Approach

INTRODUCTION

Although the detrimental life cycle consequences of teen motherhood have been well documented (An, Haveman, and Wolfe 1993; Hofferth and Hayes 1987), it is still theoretically unclear whether the negative outcomes among teen mothers result from the incidence of childbearing *per se* or from the socioeconomic disadvantages they faced during childhood and adolescence. While human capital theory holds that teenage childbearing has an exogenous effect on its socioeconomic outcomes given that it directly interferes with adolescent investment in human capital (Becker 1993), the revisionist view contends that teenage childbearing has an endogenous effect because it occurs mostly among disadvantaged female adolescents (Geronimus, Korenman, and Hillemeier 1994). Isolating the effect of teen motherhood creates a considerable methodological challenge known as selection bias (Winship and Mare 1992; Winship and Morgan 1999). If both observed and unobserved preexisting characteristics of teen mothers account for the relationship between teenage childbearing and its socioeconomic consequences, any assertion of its causality becomes vulnerable.

There has been a heated 30-year debate about the causal role of teenage childbearing, and yet finding its “true” effects has been elusive (Hoffman 1998; Ribar 1999; Wu and Wolfe 2001). In the presence of selection bias, social scientists have developed enhanced alternative models (Cherlin 2001; Korenman, Kaestner, and Joyce 2001). Since analyses based on standard regression methods are not likely to be robust due to their failure to adequately control for preexisting socioeconomic differences between teen mothers and non-teen mothers, most of the alternative models have concentrated on finding better comparison groups. For example, within-family fixed-effects models are designed to control for unobserved family-level heterogeneity by comparing teen mothers with their sisters who gave birth after their teenage years (Geronimus and Korenman 1992). Quasi-natural experimental approaches attempt to take the approximate randomization procedures with observational data, treating twin births or miscarriages as comparison cases (Grogger and Bonars 1993; Hotz, McElroy, and Sanders 1997). Instrumental variables methods utilize variables that capture the exogenous component of teenage childbearing to mitigate the selection bias problem (Klepinger, Lundberg, and Plotnick 1995). Even if intuitively appealing, all of these models have their own drawbacks. As discussed below, it is not uncommon to find that they are grounded on somewhat strong assumptions and/or unrepresentative samples.

In this paper, I propose a propensity score matching approach to improve on identification of the early socioeconomic effects of teenage childbearing. Following the counterfactual framework (Rosenbaum and Rubin 1983; Rubin 1977), this approach matches teen mothers (“treatment” group) to female adolescents who did not give birth as teens but are similar in all other preexisting observed characteristics (“control” group), based on a propensity to give birth as teens. Then it compares various socioeconomic outcomes between these two groups using semi- and non-parametric estimators. I seek to extend the previous literature in three distinct ways. First, this study addresses selection bias due to both observed and unobserved covariates in obtaining matching estimates of the effects of teen motherhood. Although some research has employed the matching approach to examine differences in educational attainment between teen mothers and non-teen mothers, it focuses mainly on selection bias due to observed covariates (Chevalier and Viitanen 2003; Levine and Painter 2003). Using a sensitivity analysis developed by Rosenbaum (2002), this study evaluates the magnitude of hidden bias that may nullify the causal effects of teenage childbearing estimated from propensity score matching. Second, I assess differences in labor market and welfare outcomes as well as educational attainment between teen mothers and non-teen mothers to provide a more comprehensive picture of the early socioeconomic consequences of teen motherhood.

Third, this study uses data from the National Longitudinal Study of Adolescent Health (Add Health), which cover the latest cohort of adolescents who transition into young adulthood with a broader age range (adolescents in grade 7 to 12 in 1995) and contain a variety of multilevel covariates that have not been incorporated in previous research. This paper clearly recognizes that any estimates of the socioeconomic effects of teenage childbearing should be put in context. The teen birth rate in the United States has steadily declined since 1991, when it was at its peak (Martin et al. 2006).[1] In addition, over the past 25 years, the economic returns to education have increased, a new welfare policy—Personal Responsibility and Work Opportunity Act (PRWORA)—introduced in 1996 replaced Aid to Families with Dependent Children (AFDC) with Temporary Assistance for Needy Families (TANF), and the influx of immigrants, especially Hispanics, has continued to grow. While the implications of all of these changes on the direction and magnitude of the socioeconomic effects of teenage childbearing are unclear, Add Health can help to address the possible impact of the recent contextual changes. Taken together, the counterfactual approach taken here is expected to provide a useful insight into estimating the causal effects of teenage childbearing.

SOCIOECONOMIC EFFECTS OF TEENAGE CHILDBEARING: IS IT CAUSAL?

Theoretical Views

[1] Nonetheless, the teen birth rate in the U.S. has still been the highest among the developed countries (Ventura et al. 2004).

Human capital theory argues that teenage childbearing may have deleterious consequences (Becker 1993). According to this theory, the incidence of early childbearing tends to raise the opportunity costs of accumulation in human capital. Being a mother during adolescence may hinder human capital investment because it is during this critical period that one's education is attained. Given high secondary school dropout rates of teen mothers, they are less likely to attain a college degree, which is more valued in labor markets.[2] In addition, teenage motherhood may keep young mothers from participating in the labor force due to their low educational attainment and to the incompatibility between employment and child rearing. As a result, teen mothers tend to be more dependent on welfare and trapped in poverty. This point of view places more emphasis on the detrimental effects of teenage childbearing rather than the disadvantaged origins of these adolescent women that could have led to teen motherhood. Although teen mothers mostly come from disadvantaged families, giving birth as a teen further lowers the chance for a woman to escape prolonged poverty (Furstenberg 1991). It is argued that since young mothers are still at an early developmental stage of life, they are unlikely to take the appropriate economic, social, and psychological responsibilities for their own attainment and child rearing.

In opposition to this conventional view, the revisionist view maintains that teenage childbearing does not necessarily cause negative consequences for young mothers (Geronimus 1991). Given that the majority of teen mothers come from impoverished families and neighborhoods, simply postponing childbearing during adolescence may not be sufficient for disadvantaged young women to escape poverty. Furthermore, faced with poor conditions and bleak prospects, they may have an incentive for early childbearing as an adaptive

[2] The General Educational Development (GED) has been another route to postsecondary education. Upon dropping out of high school, teen mothers are no less likely than other female high school dropouts to obtain the GED (Upchurch and McCarthy 1990). But the GED's intended effect is unclear: male GED recipients are akin to high school dropouts rather than high school graduates (Cameron and Heckman 1993), whereas female GED recipients fare better than high school dropouts but worse than high school graduates (Cao, Stromsdorfer, and Weeks 1996).

strategy. It can be regarded as a culturally rational response to poverty because childbearing could derive socioeconomic support from extended families and neighborhoods (Geronimus, Korenman, and Hillemeier 1994). From this revisionist point of view, teen mothers are considered disadvantaged young women attempting to increase economic prospects adapting to negative conditions of poverty, rather than adolescents suffering from a deleterious event of early child bearing. Since the adverse consequences of teenage childbearing may be an artifact of the preexisting socioeconomic disadvantages faced by teen mothers, this view emphasizes substantive knowledge about the ways in which teen mothers might systematically differ from other young women in order to assess the causal effects of teenage childbearing.

Alternative Models

These two contrasting views raise an important empirical question as to how one can take omitted variables and selection biases into account. To the extent that teen mothers differ systematically from other young women with respect to preexisting characteristics, the causal relationship between teenage childbearing and its socioeconomic outcomes cannot be established. For this reason, a variety of novel approaches have been taken to clarify the causal link between teenage childbearing and its socioeconomic consequences (see Hoffman (1998) for a comprehensive review). These alternative models carefully attempt to account for not only observed but also unobserved differences between teen mothers and other young women. They include within-family fixed-effect models, instrumental variables methods, and quasi-natural experimental models.

Within-family fixed-effect models compare sisters whose childbearing was timed at different ages. Since sisters share the same family and neighborhood characteristics, comparing sisters is expected to eliminate unmeasured environmental factors. Geronimus and Korenman (1992, 1993) show that cross-sectional studies overstate the correlation between teenage childbearing and its negative socioeconomic outcomes, and the effects of early motherhood are minimal in most cases. However, the sister comparison bears several substantive concerns (Hoffman, Foster, and Furstenberg 1993a, 1993b). Its estimation is based on somewhat small samples in which sisters co-resided at the time of survey. Even comparing teen mothers with their sisters who formed their own households does not strengthen the representativeness of the sample, because the sister study requires a sample from relatively large families. More importantly, within-family fixed-effects models are not very powerful in capturing individual differences—especially time-varying—between teen mothers and their sisters. If teen mothers have lower levels of cognitive and noncognitive abilities than their sisters, a failure to control for these attributes will make the negative effects of teenage childbearing biased upward (e.g., Holmlund 2005); on the other hand, if female adolescents give birth as a strategic response to poverty and their sisters, who have grim socioeconomic prospects, are more likely to co-reside at home, the negative effects of teenage childbearing will be biased downward. In addition, the assumption of shared (i.e., fixed) family background factors is violated if parents divert resources away from the non-childbearing sister to the sister who has a birth, also biasing the negative effects downward.

Instrumental variables methods are designed to take the endogeneity of teenage childbearing into account, with most attention paid to finding variables that must be correlated with teenage childbearing but otherwise unrelated to its socioeconomic outcomes. For instance, age at menarche and a variety of measures regarding abortion have been utilized as instrumental variables (Klepinger, Lundberg, and Plotnick 1999; Olsen and Farkas 1989; Ribar 1994). Although all these variables are justified in a statistical sense, the assumption that they have no relation to young women's socioeconomic outcomes is untestable.[3]

[3] Moreover, in most cases, the instrumental variables methods only estimate the local average treatment effect (LATE) (Angrist and Krueger 2001; Imbens and Angrist 1994). These methods

Bringing the first two together, quasi-natural experimental approaches attempt to identify better comparison groups to teen mothers, such as teen mothers who had twin births or pregnant adolescents who had miscarriages. These comparison groups would not differ from other teen mothers in characteristics, because twin births and miscarriages arguably occur randomly. Twin studies show that teenage childbearing has modest but adverse effects on women's subsequent outcomes for blacks and to a lesser degree, for whites (Grogger and Bonars 1993), while miscarriage studies find that most of the negative effects of teenage childbearing are short-lived and its effects become positive when teen mothers reach their mid- and late 20s (Hotz, Mullin, and Sanders 1997; Hotz, McElroy, and Sanders 2005). However, the findings from both studies need to be cautiously interpreted because twin births and miscarriages are rare events. In addition, teen mothers with twins might benefit from economies of scale compared with teen mothers with one child (Hoffman 1998). This would result in an underestimation of the effects of teenage childbearing. Similarly, miscarriage studies would underestimate its effects if the underreporting of miscarriages and/or abortions removes young women whose early pregnancies are stigmatized from the comparison group.

A COUNTERFACTUAL APPROACH: PROPENSITY SCORE MATCHING

A Matching Framework for Causal Inference

As depicted in the alternative models above, an assessment of the "true" effects of teenage childbearing on socioeconomic consequences involves the fundamental problem of causal inference: One cannot simultaneously observe the outcome of interest because female adolescent cannot be a teen mother (being a treated subject) and a non-teen mother (being a control subject) at the same time (Holland 1986). In the experimental setup, this problem may be solved by randomization because it would ensure that the treatment group is identical

---

would provide an estimate for only a subset of female teenagers who changed their childbearing behavior due to the early initiation of menarche or the availability of state abortion facilities.

to the control group in all characteristics other than the treatment assignment. Any remaining differences in the outcome between the two groups, thus, are attributable to the causal effect of the treatment. In most social scientific studies, however, random assignment is not feasible. Teenage childbearing occurs *nonrandomly* as it is concentrated among the disadvantaged subpopulation. This selection bias problem invokes an important but often neglected issue: one should estimate the causal effect of teenage childbearing on young women in the sample who experienced teenage childbearing rather than on young women as a whole.[4] To do so, it is crucial to identify a reliable comparison group that does not experience teen motherhood but is similar to teen mothers in their preexisting characteristics.

In the spirit of the counterfactual analysis with observational data, this study takes a propensity score matching approach.[5] Basically, one could imagine that two young women are matched on the same preexisting observed characteristics, one of whom is a teen mother and the other is not. Each unit of observation could be stratified into subgroups according to a specific value of covariates. As more covariates are needed to ensure the determinants of teenage childbearing, however, there should be an increasing number of cells that contain no comparison unit.[6] Rosenbaum and Rubin (1983) suggest that propensity score matching greatly reduces the high dimensionality of the observed covariates by constructing a propensity score, which is defined as the probability of receiving treatment assignment. If the

[4] In the literature, this statement means that as far as the selection bias problem is concerned, the causal effect is assessed by estimating the average treatment effect for the treated (ATT) rather than the average treatment effect for the entire population (ATE) (Dehejia and Wahba 2002; Harding 2003). As discussed below, traditional regression analysis also can estimate ATT, but only with strong assumptions.

[5] For a formal exposition of this approach, see Dehejia and Wahba (2002) and Morgan and Winship (2007). Other sociological research that takes this approach includes the educational effect of Catholic school attendance (Morgan 2001), the adolescent outcomes of neighborhood child poverty (Harding 2003), and the effect of adolescent first sex on mental health (Meier 2007).

[6] Even if all n covariates are binary, the number of possible values for the covariates will increase exponentially, resulting in 2^n (Dehejia and Wahba 2002). For instance, if 10 binary variables have an effect on the incidence of teenage childbearing, there will be 1024 cells, each of which has to have at least one teen mother as well as one non-teen mother.

true propensity score is known, a pair of the treatment and the control groups matched by their true propensity scores would, in expectation, be balanced on both observed and unobserved preexisting characteristics. Since this is improbable in practice, the predicted probabilities of receiving treatment assignment obtained from a logit (or probit) model serve as estimated propensity scores that are used for matching the treated and the control groups based on preexisting observed covariates.

This approach takes four steps. First, I calculate the predicted probabilities of being a teen mother with a logit model. In this model, all observed covariates are measured prior to occurrence of teenage childbearing. I use these predicted probabilities as the propensity scores. Drawing on previous literature, I select a variety of sociodemographic, individual, school, and neighborhood characteristics as the determinants of the incidence of teenage childbearing (see the data and measures section below).

Second, a sample that consists of teen mothers and matched non-teen mothers is generated using the propensity scores. Among non-teen mothers, only those who are close enough to teen mothers in terms of the propensity scores are included in this matched sample. Among a variety of matching algorithms, this study considers single-nearest-neighbor caliper matching with replacement (Morgan and Harding 2006).[7] The caliper size set in this analysis is .01, which restricts matches to have differences in the propensity scores within 1 percentage point.

Third, I examine whether matched teen mothers and non-teen mothers are balanced on observed covariates. I test whether the logit model from which the propensity scores are calculated achieves a significant reduction in absolute bias—the standardized percentage mean difference in each covariate between the treatment and control groups.[8] If this

[7] Switching the matching algorithms does not alter the findings shown here (the results are available upon request from the author).

propensity score estimation model is well specified, there should be little difference between these two groups in terms of preexisting observed covariates.[9]

Fourth, I assess differences in means on the socioeconomic outcomes between matched teen mothers and non-teen mothers. To minimize reverse causality, all outcome variables are measured after teen birth. For the purpose of comparison, I also present standard regression estimates of the socioeconomic effects of teenage childbearing in the results section.

It should be noted that the propensity matching approach would produce similar estimates to those of regression-based models if these models met all of the prerequisite assumptions and had well-supported data. However, when these conditions cannot be satisfied, the propensity score matching method has clear advantages over previous studies taking parametric approaches. First, it avoids serious mismatches between those who gave teen birth and those who were least likely to give teen birth, by matching only similar cases. Standard regression methods are likely to retain these mismatches, leading to producing unrealistic average treatment effect. Second, the propensity score matching approach is semi- and non-parametric with no assumption about functional forms. In contrast, standard regression methods should assume a specific functional form, which tends to make extrapolation and smoothing inevitable in order to alleviate biases due to serious mismatches. Third, the propensity score matching estimators are known to be more efficient and free from collinearity because only the estimated propensity scores are required.

Lastly, and most importantly, this approach facilitates a relatively simple assessment of the sensitivity of matching estimates to unobserved heterogeneity. It should be noted that

---

[8] The absolute bias is calculated as a percentage of the mean difference divided by the average standard deviation for the two groups:

$$\text{Bias} = \frac{\left|100(\overline{X}_T - \overline{X}_C)\right|}{\sqrt{(s_T^2 + s_C^2)/2}}.$$

[9] Note that to achieve full optimal balance, the entire joint distribution of the matching variables must be the same. See Diamond and Sekhon (2005) for this issue in more detail.

both regression and propensity score matching methods assume no selection bias due to unobserved covariates, which posits that conditional on preexisting observed covariates, teenage childbearing is independent of the outcome of interest (Rubin 1977). However, the matching framework used here can evaluate this *ignorable treatment assignment* assumption by incorporating a sensitivity analysis, which addresses how large the selection bias problem should be to completely wipe out propensity score matching estimates of the causal effect of teenage childbearing.

Despite these advantages, there are several concerns with the propensity score matching approach. First, because some non-teen mothers whose propensity scores are distant from those of teen mothers are discarded, a certain portion of teen mothers still would not have their counterfactuals in the matched cases. In this case, one can only estimate the causal effects of teenage childbearing for the subset of the treated group that overlaps with its comparison group (Heckman, Ichimura, and Todd 1998). This *common support* problem, however, can also shed light into how comparable teen mothers and non-teen mothers are to each other in terms of preexisting observed characteristics, a topic not dealt with in previous research. Second, treatment effects evaluated in the propensity score matching framework are binary in most cases, even though a number of treatment variables of interest may be multinomial or continuous. Forcing these variables to be dichotomous is likely to cause the loss of information and statistical power (see Imbens (2000) and Kosuke and van Dyk (2004) for recent development). However, this should not be a problem here because teenage childbearing, the treatment variable in this analysis, is by definition binary.

In sum, the propensity score matching method provides a more rigorous examination of the causal effect of teenage childbearing, by relaxing the assumptions regression-based methods make and evaluating its estimates in the presence of selection bias. In the next section, I demonstrate how a sensitivity analysis can be employed in the matching framework to address the selection bias problem.

A Sensitivity Analysis: Use of the Rosenbaum Bounds

The counterfactual analysis of teenage childbearing may be sensitive to "hidden bias" due to preexisting unobserved characteristics that influence both the incidence of teenage childbearing and its socioeconomic outcomes, even if this approach achieves a balance between matched teen mothers and non-teen mothers in terms of preexisting observed characteristics. Previous studies taking the propensity matching approach do not fully account for this possibility (Chevalier and Viitanen 2003; Levine and Painter 2003). For example, a researcher might not know how families decide where to live. If this decision-making process matters both for the incidence of teenage childbearing and for its subsequent outcomes, the true effects of early motherhood will not be estimated correctly. In order to evaluate the causal effects estimated from the propensity score matching analysis, the sensitivity analysis developed by Rosenbaum (2002) addresses the strength of such an unobserved variable that would relate both to being a teen mother and to resulting in a particular socioeconomic status (see Appendix A for a formal notation).

The Rosenbaum bounds method of sensitivity analysis assumes that a confounding unobserved covariate, U, exists that affects the odds of being assigned to the treatment, T, conditional on observed covariates, X. If U has nothing to do with T, then the assignment process is regarded as random. But as the influence of U on T becomes stronger, the confidence interval on the estimated effect of T becomes wider, and the significance level of the test of the null hypothesis of no effect of T on the outcome increases (i.e., the p-value goes up). In this scenario, one gauges the end points on the bounds for the significance level of the test of the null hypothesis for each assumed level of association between U and T. This enables one to find the case where the effect of U on the outcome is so strong that knowledge about U would almost perfectly predict the level of the outcome, whether or not a unit of observation received treatment assignment. In this light, while not calibrating the exact size of the true effect of teenage childbearing in the presence of unobserved heterogeneity, the Rosenbaum bounds method provides a basis for assessing this selection bias problem by making explicit the extent to which the ignorable treatment assignment assumption underlying the propensity score matching is vulnerable (DiPrete and Gangl 2004).

A key feature of this method is to allow a researcher to benchmark the strength of unobserved confounding variables against observed variables, given that many of the determinants of the incidence of teenage childbearing have been identified in the literature. For instance, family structure is known to have a powerful effect on early motherhood and its socioeconomic consequences (Wu and Martinson 1993). By comparing the magnitude of hidden bias with that of its equivalents obtained from the known observed covariates, we can examine the strength of selection bias due to unobserved covariates required to alter the causal inference about the effects of teenage childbearing on socioeconomic outcomes.

DATA AND MEASURES

Data

This study uses data from the National Longitudinal Study of Adolescent Health (Add Health). Add Health is a nationally representative, school-based, longitudinal study of adolescents in grade 7 to 12 in 1994-1995 (see Harris et al. (2003) for more information). An In-school questionnaire was administered to more than 90,000 adolescents who attended each selected school on a particular day during the period of September 1994 to April 1995. Based on the school rosters, a random sample of about 200 students from each high school and feeder school pair was collected between April and December 1995 to yield the core In-home sample of about 12,000 adolescents. The In-home interviews added special over-samples that included racial/ethnic minorities, physically disabled adolescents, and a genetic sample. These Wave I data produced a total sample size of 20,745 adolescents, 10,480 of which are female.[10] Their parents also were interviewed in Wave I. In 2001 and 2002, approximately 15,200 Wave I respondents, 8,030 of which are female, were re-interviewed in Wave III to investigate the influences that experiences in adolescence have on young adulthood.[11]

Because of its strong emphasis on social contexts such as families, schools, and neighborhoods, and its broad definition of health-related behaviors, Add Health provides valuable information well suited for this study. First, the Wave III sample contains data on fertility, educational attainment, labor market performance, and welfare receipt. Second, the Wave I sample provides rich sets of multilevel variables that measure observed covariates prior to the incidence of teenage childbearing, which are found to affect not only early motherhood but also adolescent mothers' socioeconomic outcomes.[12] Among these potential preexisting characteristics are differences in family, school, and neighborhood environment as well as individual differences in cognitive ability and attitudes and behaviors. Third, the Wave I sample contains variables that are previously unmeasured but considered key factors of teenage childbearing and various socioeconomic outcomes. For example, recent evidence shows that noncognitive skills such as "soft skills" play an important role in teenage

[10] On the surface, it might appear that Add Health is not representative of all adolescents because it uses a school-based design and misses high school dropouts in the In-school survey. It should be noted, first, that annual dropout rates are very low at the national level (Udry and Chantala 2003). Second, the In-home sample on which this analysis is based was selected from the school rosters, which were collected approximately a year before the Wave I In-home interviews were selected. That is, the only dropouts missed were those who dropped out of school two years prior to the Wave I In-home interviews. Therefore, although missing dropouts remains a concern, it appears to be minor (see Udry and Chantala (2003) for more details).

[11] The age range at Wave III is from 18 to 27 with the average age 22. The fact that the data used here consist of young adults implies that some of the outcome variables—labor market and welfare outcomes in particular—may not be appropriate as longer-run socioeconomic consequences of teenage childbearing. Despite the strong correlations between early and late socioeconomic behaviors, the results reported in this paper refer only to the *early* socioeconomic effects of teen motherhood.

[12] There were older teenagers who were at risk for giving birth prior to Wave I. Since all covariates come from the Wave I sample, this might cause a reverse causality problem between teenage childbearing and some of the covariates. I reran all models presented here with and without those who became mothers by Wave I. In addition, all of these models were reanalyzed by limiting the analytic sample to adolescents who did not give birth and were younger than age 16 at Wave I. The results in both cases (available upon request from the author) do not alter the findings shown in this paper.

childbearing and subsequent socioeconomic outcomes, but the previous literature has not taken those measures into account (Heckman, Stixrud, and Urzua 2006; Plotnick 1992).

Finally, Add Health allows me to explore the socioeconomic effects of teenage childbearing in the late 1990s and the early 2000s, whereas past research has been limited to examining its effects only as recently as the early 1990s (e.g., Levine and Painter 2003). The issue of timeliness is important given the significant social contextual changes since that time (Hoffman 1998). The 1990s witnessed growing economic return to education, changes in welfare policy, and the increase of Hispanics in the U.S. adolescent population, all of which could have influences on adolescents' fertility behavior. This study does not address the direct impacts of these social changes on the association between teenage childbearing and its socioeconomic consequences, but does discuss the implications of recent contextual changes for the socioeconomic effects of adolescent fertility.

The analytic sample for this study consists of 6,825 adolescent females who were interviewed at Waves I and III and who had observations on teenage childbearing status, the independent variables of interest, and sampling weights. This study reports descriptive results based on the weighted sample and multivariate results based on the unweighted sample. Preliminary analyses (available upon request from the author) suggest little difference between the weighted and unweighted samples. To account for the sampling design effects in Add Health, all analyses adjust standard errors for school-level clustering.[13]

Measures

DEPENDENT VARIABLES

[13] Taking school-level clustering into account is similar in spirit to Levine and Painter (2003)'s within-school propensity score matching method. Although it is plausible that this method reduces selection bias resulting from school-level sorting, it cannot rule out selection bias involving individual- and family-level unobserved characteristics. This paper employs the Rosenbaum bounds method to adjudicate the strength of selection bias as a whole.

As one of the key dependent variables, educational attainment is a good proxy for young women's future socioeconomic status. This study evaluates differences between teen mothers and non-teen mothers with two measures of educational attainment: dropping out of high school and attending or graduating from some college (2 year or more). I treat GED recipients as high school dropouts. Since there is little consensus on whether GED recipients are considered high school dropouts or graduates (Cameron and Heckman 1993; Heckman and Rubinstein 2001; Upchurch and McCarthy 1990), I rerun the analyses to treat GED recipients as high school graduates. Also, I use attending or graduating from college (4 year or more) instead of some college for another robustness check. The results (available upon request from the author) are almost identical to those shown below.

As a second set of dependent variables, teen mothers' labor market performance relative to that of non-teen mothers is measured with employment status, work-related activities, full-time/part-time status, and weekly wages. All these variables are measured at the time of the survey at Wave III. Work-related activities include on-the-job-training as well as employment statuses. Weekly wages below the 3^{rd} percentile in the wage distribution are set equal to the 3^{rd} percentile, and weekly wages above the 97^{th} percentile are set equal to the 97^{th} percentile. Then log-transformed weekly wages are used for analyses. There are respondents who still were enrolled in post-secondary schools at Wave III, so I restrict the samples to respondents who were not in these schools. I further restrict the samples to respondents who were employed at Wave III for full-time/part-time status and those who worked full-time for weekly wages.

As a third dependent variable, welfare receipt signifies an important dimension of young women's socioeconomic consequences. This study measures welfare receipt by whether respondents were on welfare at Wave III and ever on welfare. A respondent is identified as welfare-dependent if she received AFDC (or TANF), public assistance, welfare payments, or food stamps. Due to eligibility, I restrict the sample to respondents who gave birth regardless of teenage childbearing status.

EXPLANATORY VARIABLES

I construct measures of teenage childbearing from the Add Health life history calendar of the Wave III sample. I also utilize the Add Health household rosters because not all teen mothers are identified in the life history calendar. Most previous studies of adolescent fertility and its socioeconomic consequences offer little rationale on which age cutoff point (e.g., under age 18 or 20) is used. I treat a woman as a teen mother if she gave birth prior to age 18 for educational attainment outcomes because most adolescents complete their high school education by age 18. For labor market and public assistance receipt outcomes, I treat a woman as a teen mother if she gave birth prior to age 20, given the prolongation of the stages of adolescence that has been observed over the last few decades in the U.S. As traditionally defined adult roles have been assumed at later ages, there has been a lengthening of the adolescent transition to adulthood (Rindfuss 1991).[14] Both measures of teenage childbearing used in this paper do not differentiate whether teen mothers were married at the time of childbearing, but the results reported here do not change when I exclude married teen mothers from the analytic sample (these results are available upon request from the author).

For preexisting control variables, a set of demographic and family characteristics measured at Wave I are included to be used as controls for modeling the socioeconomic effects of teenage childbearing. This study measures age, race/ethnicity, and immigrant generation status for demographic characteristics. Race/ethnicity is classified as non-Hispanic whites, non-Hispanic blacks, Hispanics, and Asians. Immigrant generation status is

[14] A supplemental analysis using alternative age cutoff points (age 20 for educational attainment outcomes and age 18 for all other outcomes) shows that changing age cutoff points does not alter the findings reported here (not shown). Note that except for the educational attainment outcomes, tables and figures in the results section are based on the analysis using an age cutoff point of 20 and they are substantively the same as those obtained from using age 18 as a cutoff point.

defined as 1st generation if a respondent is foreign-born to foreign-born parents, 2nd generation if she is U.S.-born to foreign-born parents, and 3rd or higher generation if she is U.S.-born to U.S.-born parents. Family background covers family structure, parental education, and number of siblings. Family structure is categorized as two-biological parent families, two-parent step families, single-mother families, single-father families, and other families (e.g., foster families). Parental education is measured with the highest level of education either of the parents obtained and categorized as less than high school, high school graduation, some college experience, and college graduation or more, with an indicator of missing observations on parental education.

To capture individual-level differences between teen mothers and non-teen mothers, this study measures parental monitoring, cognitive and noncognitive skills, religiosity, and risk behaviors. Parental monitoring is measured by calculating the total count of their activities monitored by parents, including curfews, friendships, TV watching, and food and dress choices. For a measure of cognitive skills, I use the Add Health Picture Vocabulary Test (AHPVT) and grade point average (GPA). AHPVT is an abbreviated version of the Peabody Picture Vocabulary Test with age-standardized scores for adolescents. I retain the missing cases on AHPVT by assigning these cases to the sample mean value and including a missing data indicator variable in the analysis. This study improves on the literature by including measures of noncognitive skills, which are referred to as attitudinal and behavioral personal traits that are correlated with but distinct from cognitive skills. The Add Health Wave I sample provides 9 items that are included in either the Rotter's locus of control scale or the Rosenberg's self-esteem scale.[15] Based on these items, I construct a composite measure of

[15] Locus of control measures the degree of control individuals feel ranging from external to internal. According to Rotter (1966), individuals who believe that outcomes are due to luck have an external locus of control while individuals who believe that outcomes are due to their own efforts have an internal locus of control. The self-esteem scale measures perceptions of self worth (Rosenberg 1965). Both scales have been commonly employed in past research on the effects of noncognitive skills on socioeconomic outcomes (Heckman, Stixrud, and Urzua 2006).

noncognitive skills with 4-point Likert scale (α = .67). Religiosity is a composite measure with 4-point Likert scale (α = .86) of attendance to religious services (from once a week or more to never), the importance of religion (from very important to not important at all), and the oftenness of prayer (from at least once a day to never). Risk behaviors are measured with the questions of whether a respondent smoked regularly and how many days per month a respondent drank alcohol during the past 12 months.

School facilitates interactions of adolescents with teachers and peers by way of providing role models and developing adaptive strategies. I focus on measures of collective socialization (Coleman 1990). These include 1) the school's structural characteristics, such as the percentage of white students in school, school type (private/public), and school region and 2) school climate, such as school-level expectations of going to college and earning a middle-class income by age 30. Also, socioeconomic conditions of neighborhood might define an individual's opportunity structure and the normative climate during adolescence and subsequently affect their future outcomes (Massey and Denton 1993; Wilson 1987). Several measures of census tract-level characteristics are constructed including urbanity, the percent idle, defined as the percentage of young people who were neither at work nor in school, and total unemployment rate.

RESULTS

Preliminary Results

Table 4.1 presents descriptive statistics by teenage childbearing status. There are 1,266 (or 18.5%) teen mothers and 5,559 (or 81.5%) non-teen mothers in the full sample. It shows statistically significant mean differences between these two groups in 25 out of 36 characteristics that were measured prior to the incidence of teenage childbearing. For demographic and family characteristics, compared to their counterparts, teen mothers were more likely to be black or Hispanic, more likely to come from the 3rd+ generation, less likely to reside with two-biological parents, and less likely to have parents with college diploma. For individual characteristics, teen mothers had lower cognitive and noncognitive skills and were more likely to smoke and drink alcohol than non-teen mothers. The schools that teen mothers attended had more minorities, were more likely to be public, had lower levels of group expectations of going to college and earning a middle-class income, and were more likely to be in the South than the schools that their counterparts attended. Lastly, teen mothers were more likely than non-teen mothers to reside in neighborhoods with higher percentages of the idle and total unemployment.

<< Table 4.1 is about here. >>

Logit estimates that predict teenage childbearing status by all covariates also confirm the overall picture depicted above. Table 4.2 shows that compared to young women who did not give teen birth, teen mothers tend to come from socioeconomically disadvantaged populations. Teen mothers are statistically different from their counterparts at the .05 level in terms of almost all preexisting covariates. These differences raise a critical question as to how one can find a better comparison group to teen mothers in order to estimate the causal effects of teenage childbearing on socioeconomic outcomes. In Figure 4.1, I compare teen mothers and non-teen mothers in the full sample by way of their predicted probabilities of teenage childbearing that are obtained from the logit model in Table 4.2. Clearly, there is much discrepancy between the two groups: serious mismatches exist between teen mothers and non-teen mothers in the region where a predicted probability to give teen birth is higher. Recall that standard regression methods should impose strong assumptions to control for potential biases resulting from those mismatches. Propensity score matching estimates will show how much biases regression-based estimates could produce in this respect.

<< Table 4.2 is about here. >>

<< Figure 4.1 is about here. >>

Matching Results

Following the matching framework, I construct a matched sample based on the estimated propensity scores of teenage childbearing. The matched sample consists of 1,259 teen mothers and 1,259 non-teen mothers whose propensity scores are sufficiently close to those of teen mothers. Figure 4.2 depicts the degree to which teen mothers and non-teen mothers overlap with each other in the matched sample. Strikingly, these two groups are well matched, indicating that there is little difference in the preexisting observed socioeconomic covariates.[16]

<< Figure 4.2 is about here. >>

Table 4.3 presents propensity score matching results of the causal effects of teenage childbearing on socioeconomic consequences. The first and second columns of each panel report the standard regression and propensity score matching estimates, respectively, for each outcome. The regression estimates are obtained from logit or OLS models using teenage childbearing status and all covariates in Table 4.2 as the independent variables. They are expressed as the simulated predicted probabilities that are computed by averaging each respondent's value on all covariates except for teenage childbearing status. The matching estimates are the simple mean probabilities of each outcome in the matched sample. To make these two estimates comparable, the ratios between the treatment and control groups for each outcome are presented in the fourth rows of each panel. As indicated in the last rows of each panel, all outcome-specific matched samples have strong common support, ranging from 99.2% to 99.8%. This indicates that the propensity score matching method succeeds in

[16] In the Appendix B, Table B1 gives an additional snapshot of the covariate balance check. None of the preexisting covariates bears statistical differences in means between teen mothers and matched non-teen mothers; the matched sample also achieves a significant percentage reduction in absolute bias. Although three variables—father only family, parents with some college education, and Midwest—show an increase in absolute bias (-22.3, -7710.3, and -173.4, respectively), these variables do not statistically differ between teen mothers and non-teen mothers before as well as after matching. Note that using the outcome-specific matched samples produces the almost same pattern of overlapping and covariate balance between these two groups (not shown).

locating a sufficiently large number of non-teen mothers who share similar preexisting observed characteristics with teen mothers.

<< Table 4.3 is about here. >>

With respect to educational attainment, Panel A1 shows the propensity score matching estimate of the effect of teen motherhood on dropping out of high school, suggesting that when taking the ratio of the dropout rate between teen mothers and non-teen mothers, teen mothers are about 2.2 times more likely than non-teen mothers to be dropouts (.339/.154=2.205). This estimate is less than half the logit estimate (=5.692) but still statistically significant. Panel A2 shows that teen mothers are about 40% less likely to attend or graduate from some college than their matched counterparts (1–.585=.415), which is an estimate that is much smaller than the logit estimate (1–.310=.690) but remains statistically significant. For labor market outcomes, the propensity score matching estimates also find the statistically significant negative effects of teenage childbearing, although the magnitudes of its effects reduce compared to the logit estimates. Panels B1 through B3 show that compared to matched non-teen mothers, teen mothers are about 15% less likely to be employed (1–.845=.155), about 12% less likely to participate in work-related activities (1–.883=.117), and about 9% less likely to work full-time (1–.906=.094). The logit estimates for each of these labor market outcomes are about 25% (1–.753=.247), 18% (1–.821=.179), and 9% (1–.905=.095), respectively.

Teen mothers do not differ from their matched counterparts in terms of weekly wages and public assistance receipt, as shown in Panels B4, C1 and C2. Note, however, that for weekly wages, the analytic sample only consists of those who worked full-time. This implies that teen mothers in this sample are a selected group among teen mothers as a whole. For the welfare outcomes, little difference between teen mothers and non-teen mothers is not likely to be driven by the preexisting socioeconomic disadvantages facing teen mothers, given that teen mothers perform more poorly in the labor market than do non-teen mothers. A potential hypothesis is that since the post-1996 welfare policy applies more strict criteria on eligibility and duration of public assistance receipt, teenage childbearing status would not differentiate public assistance use for socioeconomically disadvantaged women (testing this hypothesis is beyond the scope of the paper).

The results from the propensity score matching analysis of the socioeconomic consequences of teenage childbearing show that traditional regression method tends to overestimate the negative effects of teen motherhood. Obviously, teen mothers' lower levels of educational attainment and labor market performance result substantially from their disadvantaged socioeconomic background; nevertheless, there are still sizable differences between teen mothers and non-teen mothers in the key domains of subsequent outcomes considered in this study. The matching results suggest that even faced with the similarly adverse socioeconomic conditions when growing up, young women who did not give teen birth fare better than teen mothers in educational attainment and labor market outcomes. Teenage childbearing seems to impose an additional burden on already disadvantaged young women, leading to more adverse consequences for teen mothers.[17] However, the matching estimates reported here and in Levine and Painter (2003) do not take into account selection bias due to unobserved preexisting characteristics, which may produce upwardly biased estimates of the effects of teenage childbearing. A contribution of this paper is to conduct a sensitivity analysis using the Rosenbaum bounds method to address the role of unobserved heterogeneity in making a causal inference about teen motherhood and its socioeconomic consequences.

Results from the Sensitivity Analysis

[17] These results are also consistent with Levine and Painter (2003)'s findings with regard to educational attainment.

Table 4.4 presents the Rosenbaum bounds for the causal effects of teenage childbearing. Γ in the first column indicates the magnitude of selection bias due to unobserved covariate—hidden bias—that supposedly predicts the incidence of teenage childbearing, which is expressed as an odds ratio. *p*-critical in the second column denotes the *p*-value at which a matching estimate of the effect of teenage childbearing become insignificant corresponding to the given magnitude of hidden bias. I compare the magnitudes between hidden bias (Γ) and its known equivalents to give a substantive interpretation of how large hidden bias should be to wipe out the propensity score matching estimates of the causal effect of adolescent fertility. The hidden bias equivalents are selected from the observed covariates that are found to have a significant effect on being a teen mother and computed as odds ratios or effect sizes from the logit model in Table 4.2.

<< Table 4.4 is about here. >>

First, the Rosenbaum bounds suggest that the causal effect of being a teen mother on educational attainment is most robust, compared to other socioeconomic outcomes. Panel A in Table 4.4 reports that the effect of adolescent fertility on dropping out of high school becomes statistically insignificant at the .05 level (*p*-critical is .056) as Γ approaches 2.2. This means that in order to challenge the significance of the matching estimate, an unobserved covariate should cause the odds ratio of being a teen mother to differ between teen-mothers and matched non-teen mothers by a factor of 2.2. A selection bias with such magnitude is larger than the estimated net effect of being non-Hispanic black instead of being non-Hispanic white (2.101), having parents with less than a high school education instead of college diploma (1.989), receiving the lowest GPA of 1 instead of the average GPA of 2.9, or moving into a school where only 50% of the students expect to go to college instead of the average school where 78.4% of the students expect to go to college. Thus, to drive the effect of teenage childbearing on dropping out of high school to statistical insignificance, the effect of an unobserved covariate should be stronger than the estimated effects of race, parental education, GPA, or school-level expectation of going to college even after controlling for all of these variables.

With regard to the effect of teenage childbearing on attending or graduating from some college, the Rosenbaum bounds statistics indicates that Γ should be at least 1.8 (*p*-critical is .057) to completely alter this effect. The effect of an unobserved covariate should be larger than the estimated net effect of living in a step (1.535) or single-mother family (1.674) instead of a two-biological parent family, having parents with high school diploma instead of college diploma (1.770), receiving the lowest GPA of 1 instead of the average GPA of 2.9, or moving into a school where only 50% of the students expect to go to college instead of the average school where 78.4% of the students expect to go to college.

Second, the causal effects of teenage childbearing on labor market outcomes appear to be somewhat vulnerable to unobserved confounding variables, compared to its effect on educational attainment. Panel B suggests that to nullify the effects of teen motherhood on employment status, the critical value of Γ should be 1.4 (*p*-critical is .133), which is comparable to the estimated net effect of living in a single-mother family instead of a two-biological parent family (1.460), an additional 1.2 point decline in GPA, or moving into a neighborhood in which total unemployment rate is 3% higher than the neighborhood with an average total unemployment rate. The effects of teen motherhood are a little weaker on work-related activities and full-time/part-time status than that on employment status. To make those effects insignificant, the critical values of Γ should be 1.3 (*p*-criticals are .125 and .072 respectively), which are equivalent to .9 point decline in GPA or moving into a neighborhood where total unemployment rate is 2.3% higher than the neighborhood with an average total unemployment rate. As Diprete and Gangl (2004) have pointed out, however, these results are worst-case scenarios. If selection bias due to unobserved covariates had a powerful effect on being a teen mother but only a small effect on the outcomes, the effect of teenage childbearing would not disappear. Lastly, the Rosenbaum bounds for weekly wages and public assistance receipt (Panel C) imply that the effects of teenage childbearing are highly vulnerable to hidden biases, which is consistent with the matching results shown in Table 4.3 that find no effects.

In summary, the sensitivity analysis using the Rosenbaum bounds suggests that although we cannot rule out the possibility that selection bias due to unobserved covariates may underlie the matching estimates, this sort of bias would have to be substantial in order to completely eliminate the causal effects of teenage childbearing on most of the early socioeconomic outcomes considered in this study.

Racial/Ethnic Differences

Given recent changes in the racial/ethnic composition of female adolescents in the United States (i.e. the increase of Hispanics) and the conflicting findings in regard to the effects of teenage childbearing on socioeconomic consequences by race/ethnicity, an examination of racial/ethnic differences deserves attention.[18] Table 4.5 reports the propensity score matching estimates of the socioeconomic effects of teen motherhood among non-Hispanic whites, non-Hispanic blacks, and Hispanics. For non-Hispanic whites, the matching estimates show that teen motherhood has negative effects on educational attainment, employment status, and work-related activities, but not on full-time/part-time status, weekly wages, and public assistance receipt. The Rosenbaum bounds suggest that the effect of being a teen mother on educational attainment is most robust to unobserved heterogeneity, compared to its effects on employment status and work-related activities. For non-Hispanic blacks, teen motherhood has adverse consequences on educational attainment and employment status. For Hispanics, teen motherhood has a negative effect only on high school completion and although not

[18] For example, Grogger and Bronars (1993) report that most of the adverse consequences of teen motherhood are amplified among blacks, whereas Klepinger, Lundberg, and Plotnick (1999) find that the negative effects of teen motherhood are present for both whites and blacks.

significant, its effects on employment status and work-related activities are significantly different from those for non-Hispanic whites and blacks.

<< Table 4.5 is about here. >>

On the one hand, the results imply that the socioeconomic effects of teenage childbearing are detrimental to non-Hispanic whites and to a lesser degree, non-Hispanic blacks, but its effects might be short-lived for Hispanics. The findings from Hispanic teen mothers seem in favor of the revisionist view (Geronimus, Korenman, and Hillemeier 1994). While racial/ethnic differences in the effect of teen motherhood require further research, it should be noted that a large number of respondents were enrolled in post-secondary schools at Wave III (about 41% for non-Hispanic whites, 42% for non-Hispanic blacks, and 38% for Hispanics). The socioeconomic effects of teenage childbearing might be restored if these respondents entered the labor market upon graduation. This is likely the case, given the increasing economic returns to education over the past 25 years (Card 1999; Lemieux 2006). Recall that there is significant difference in education attainment between teen mothers and non-teen mothers across all racial/ethnic groups. New data that contain older age groups may help determine which interpretation can be substantiated.

DISCUSSION AND CONCLUSION

Studies of teenage childbearing and its socioeconomic consequences have always been concerned about omitted variables and selection biases that are critical to estimate the “true” effect of teenage childbearing. Despite the apparent burden of childbearing to female adolescents, the detrimental socioeconomic preconditions facing teen mothers make it extremely difficult to sort out the relationship between teen motherhood and subsequent socioeconomic consequences. While various creative approaches such as sister fixed-effects models, instrumental variables methods, and quasi-natural experiments have been provided, unrepresentative samples and untestable assumptions have hampered these approaches and yielded mixed results. Although some studies used a propensity score matching estimator to address these problems, they paid no direct attention to selection bias on unobserved characteristics (Chevalier and Viitanen 2003; Levine and Painter 2003). Furthermore, recent changes in social context and the composition of adolescent population require a new evaluation of the socioeconomic effects of teenage childbearing. The counterfactual analysis employed in this study is designed to shed new light into this line of research by 1) finding a better comparison group to teen mother with less dependence on statistical assumptions and a larger sample size; 2) employing the Rosenbaum bounds method to assess the significance level of the matching estimates of the causal effects of teenage childbearing in the presence of selection bias on unobserved covariates; and 3) providing the most up-to-date assessment of the early socioeconomic effects of teenage childbearing with Add Health data.

As in most of the previous studies taking alternative approaches, the propensity score matching results show that socioeconomic disadvantages inherent to teen mothers account for a nontrivial portion of the effects of teen motherhood on subsequent socioeconomic outcomes, suggesting that the selection bias problem results in an overestimation of its negative effects. However, when teen mothers are compared to their matched counterparts who are similar in every preexisting characteristic except for teenage childbearing status, being a teen mother still has significant negative effects on educational attainment and most of the labor market outcomes. Moreover, a sensitivity analysis employing the Rosenbaum bounds method suggests that selection bias due to unobserved covariates would have to be very powerful to alter the propensity score matching estimates of the causal effect of teenage childbearing on educational attainment and to a lesser degree, on early labor market outcomes. These findings are consistent with some of the earlier studies using within-family fixed-effects models (Hoffman, Foster, and Furstenberg 1993a, 1993b; Holmlund 2005), instrumental variables methods (Klepinger, Lundberg, and Plotnick 1999), and quasi-natural experimental approaches (Grogger and Bonars 1993). This convergence in the findings points out that identifying more reliable comparison group to teen mothers through flexible modeling assumptions and well-supported data should be given a priority in the research on the causal relationship between adolescent fertility and its socioeconomic consequences.

There are several limitations of this study that warrant mention. First, the propensity score matching method combined with the Rosenbaum bounds should be understood as an effort to make the causal inference about the effects of teenage childbearing more constructive, rather than as the solution to resolve all complex issues regarding selection bias. Second, the results reported here refer to the *total* and *average* effects of teenage childbearing for socioeconomically disadvantaged young women. We do not know yet the exact mechanisms by which teenage childbearing affects teen mothers' educational attainment and subsequent life outcomes (Moffitt 2005). Also, if the heterogeneity among teen mothers—e.g., due in part to unobserved time-varying covariates—still exists even after controlling for their socioeconomic disadvantages, teenage childbearing may have differential impacts on their multiple life domains. Third, it is likely that structural changes influence the association between teenage childbearing and subsequent outcomes, but this study is only suggestive of the potential roles of the growing economic returns to education and the post-1996 welfare policy. Future research can benefit from more comprehensive knowledge about teen mothers' life experiences, new data with older age groups, and innovative strategies and/or cohort data linking the effects of teenage childbearing to macro-level social changes.

Despite these limitations, this paper clearly indicates that teenage childbearing is likely to be a further obstacle for disadvantaged young women to overcome in order to advance in educational attainment and early labor market performance. In this light, public policy implications that can be derived from this study are two-fold. The analysis shows that adolescents' attitudes and behaviors such as noncognitive skills and smoking affect the incidence of teenage childbearing. Research suggests that these habits and traits are more malleable during childhood than cognitive skills and stabilize during adolescence (Carneiro and Heckman 2003). Policy interventions need to be made at the early life stages of childhood and pre-adolescence. This means that reducing the teen birth rate should be carried out in a broader context of socioeconomic disadvantages that exist well before late adolescence. In addition, this study consistently finds that educational attainment is the major domain of socioeconomic disadvantages facing teen mothers. Policies aimed at raising the high school completion rate and making college education more affordable would provide opportunities for teen mothers as well as disadvantaged adolescents as a whole to improve their socioeconomic viability.

Table 4.1. Descriptive Statistics by Teenage Childbearing Status: Full Sample

	All Women				Non-teen Mothers		Teen Mothers
Variable	Mean	S.E.	Min.	Max.	Mean		Mean
Demographic and family characteristics							
Age at Wave I	15.256	0.123	11	21	15.267		15.204
White (reference)	0.675	0.031	0	1	0.706	***	0.538
Black	0.167	0.023	0	1	0.139	***	0.291
Hispanic	0.115	0.017	0	1	0.106	*	0.151
Asian	0.043	0.009	0	1	0.048	***	0.019
Immigrant generation:							
1st	0.051	0.009	0	1	0.054		0.036
2nd	0.104	0.010	0	1	0.109		0.084
3rd+ (reference)	0.845	0.019	0	1	0.837	*	0.880
Two-biological parent family (reference)	0.570	0.014	0	1	0.616	***	0.367
Step family	0.167	0.007	0	1	0.155	***	0.224
Mother only family	0.205	0.010	0	1	0.180	***	0.321
Father only family	0.023	0.003	0	1	0.023		0.024
Other family structure	0.034	0.004	0	1	0.027	***	0.064
Parental education:							
Less than high school	0.120	0.010	0	1	0.105	***	0.187
High school graduate	0.311	0.011	0	1	0.290	***	0.406
Some college	0.220	0.009	0	1	0.218		0.229
College graduate (reference)	0.315	0.016	0	1	0.355	***	0.138
Number of siblings	1.403	0.038	0	12	1.396		1.431
Individual characteristics							
Parental monitoring	5.110	0.056	0	7	5.111		5.104
Add Health PVT score	100.801	0.593	16	138	101.981	***	95.500
Grade Point Average (GPA)	2.902	0.024	1	4	2.982	***	2.544
Rotter/Rosenberg scale	2.262	0.008	0	3	2.286	***	2.155
Religiosity	1.950	0.033	0	3	1.967		1.874
Regular smoking	0.200	0.012	0	1	0.171	***	0.330
Frequency of drinking	1.098	0.067	0	30	1.047	*	1.326
School characteristics							
Percent white	63.098	2.726	0	97.619	64.540	***	56.622
School type: public	0.936	0.020	0	1	0.928	**	0.974
Percent expectations:							
Going to college	78.430	0.568	50	96.094	78.651	*	77.439
Middle class income	63.109	0.389	40.865	74.415	63.318	**	62.169
West	0.168	0.014	0	1	0.175		0.135
Midwest	0.300	0.024	0	1	0.296		0.317
South	0.396	0.018	0	1	0.377	**	0.481
Northeast (reference)	0.137	0.012	0	1	0.152	***	0.066
Neighborhood characteristics							
Urbanity	0.254	0.041	0	1	0.257		0.244
Percent idle	5.330	0.344	0	48.889	5.023	***	6.709
Total unemployment rate	7.592	0.364	0	50.909	7.212	***	9.298
N	6825				5559		1266

*** $p<0.001$, ** $p<0.01$, and * $p<0.05$ (two-tailed tests).

Table 4.2. Parameter Estimates from Logit Model Predicting Teenage Childbearing Status

	Coefficient		S.E.
Demographic and family characteristics			
Age at Wave I	-0.107	**	(0.032)
White (reference)			
Black	0.634	***	(0.110)
Hispanic	0.686	***	(0.139)
Asian	0.276		(0.216)
Immigrant generation:			
1st	-1.058	***	(0.255)
2nd	-0.459	**	(0.179)
3rd+ (reference)			
Two-biological parent family (reference)			
Step family	0.529	***	(0.092)
Mother only family	0.379	***	(0.096)
Father only family	0.004		(0.186)
Other family structure	0.650	**	(0.231)
Parental education:			
Less than high school	0.717	***	(0.127)
High school graduate	0.588	***	(0.091)
Some college	0.480	***	(0.091)
College graduate (reference)			
Number of siblings	0.040		(0.035)
Individual characteristics			
Parental monitoring	0.010		(0.026)
Add Health PVT score	-0.013	***	(0.003)
Grade Point Average (GPA)	-0.409	***	(0.051)
Rotter/Rosenberg scale	-0.282	***	(0.079)
Religiosity	-0.015		(0.034)
Regular smoking	0.856	***	(0.081)
Frequency of drinking	0.005		(0.010)
School characteristics			
Percent white	-0.001		(0.002)
School type: public	0.282		(0.172)
Percent expectations:			
Going to college	-0.022	*	(0.011)
Middle class income	0.021		(0.018)
West	0.431	*	(0.215)
Midwest	0.572	**	(0.196)
South	0.788	***	(0.202)
Northeast (reference)			
Neighborhood characteristics			
Urbanity	-0.144		(0.126)
Percent idle	0.011		(0.008)
Total unemployment rate	0.019	*	(0.009)
Constant	1.562		(1.072)
-2 log pseudolikelihood		5256.063	
N		6825	

Notes: Missing indicators for parental education and Add Health PVT score are included but not shown; Robust standard errors adjusting school-level clustering in parentheses.
*** $p<0.001$, ** $p<0.01$, and * $p<0.05$ (two-tailed tests).

Table 4.3. Propensity Score Matching Estimates of the Effects of Teenage Childbearing

	Regression Estimates		Matching Estimates		Regression Estimates		Matching Estimates	
	A. Educational Attainment							
	A1. Dropout				*A2. College attendance*			
Non-teen mothers (Control)	0.052		0.154		0.589		0.411	
Teen mothers (Treatment)	0.297		0.339		0.182		0.241	
Difference (T - C)	0.245	***	0.185	***	-0.406	***	-0.170	***
Ratio (T/C)	5.692		2.205		0.310		0.585	
Treatment cases	575		572		573		569	
Control cases	6240		572		6219		569	
Percent common support			99.5%				99.3%	
	B. Labor Market Outcomes							
	B1. Employment status				*B2. Work-related activities*			
Non-teen mothers (Control)	0.760		0.675		0.778		0.717	
Teen mothers (Treatment)	0.572		0.570		0.639		0.633	
Difference (T - C)	-0.188	***	-0.105	***	-0.139	***	-0.084	**
Ratio (T/C)	0.753		0.845		0.821		0.883	
Treatment cases	997		991		996		990	
Control cases	3021		991		3018		990	
Percent common support			99.4%				99.4%	
	B3. Full-time employment				*B4. Weekly wages (logged)*			
Non-teen mothers (Control)	0.840		0.830		5.883		5.824	
Teen mothers (Treatment)	0.760		0.752		5.829		5.830	
Difference (T - C)	-0.080	***	-0.078	**	-0.054		0.006	
Ratio (T/C)	0.905		0.906		0.991		1.001	
Treatment cases	565		564		417		416	
Control cases	2239		564		1802		416	
Percent common support			99.8%				99.8%	
	C. Public Assistance Receipt							
	C1. Currently on welfare				*C2. Ever on welfare*			
Teen mothers (Control)	0.267		0.242		0.344		0.304	
Non-teen mothers (Treatment)	0.200		0.228		0.257		0.272	
Difference (C - T)	0.067		0.014		0.087	**	0.032	
Ratio (C/T)	1.335		1.063		1.339		1.118	
Treatment cases	637		632		627		622	
Control cases	1109		632		1092		622	
Percent common support			99.2%				99.2%	

Notes: For public assistance receipt outcomes, teen mothers are treated as the control group because there are more teen mothers in the sample; Statistical significance levels calculated from bootstrap standard errors for the matched sample (300 replications).

*** p<0.001, ** p<0.01, and * p<0.05 (two-tailed tests).

Table 4.4. A Sensitivity Analysis Using the Rosenbaum Bounds of the Causal Effects of Teenage Childbearing

	Γ	*p*-critical
A. Educational Attainment		
Dropout	1.9	0.005
	2.0	0.013
	2.1	0.029
	2.2	0.056
College attendance	1.5	0.002
	1.6	0.007
	1.7	0.023
	1.8	0.057
B. Labor Market Outcomes		
Employment status	1.1	<0.001
	1.2	0.003
	1.3	0.028
	1.4	0.133
Work-related activities	1.1	0.002
	1.2	0.024
	1.3	0.125
Full-time employment	1.1	0.005
	1.2	0.023
	1.3	0.072
Weekly wages (logged)	1.0	0.592
C. Public Assistance Receipt		
Currently on welfare	1.0	0.274
Ever on welfare	1.0	0.100

Notes: Γ is the odds ratio of differential treatment assignment due due to an unobserved covariate; *p*-critical from the Wilcoxon signed rank tests; *p*-critical is *p*+ for dropout, weekly wages, and public assistance receipt outcomes, and *p*- for college attendance and other labor market outcomes.

Table 4.5. Propensity Score Matching Estimates of the Effects of Teenage Childbearing, by Race/Ethnicity

	White		Black		Hispanic	
A. Educational Attainment						
A1. Dropout	0.174	***	0.173	***[a]	0.200	**
	(0.049)		(0.037)		(0.074)	
Matched pair	218		226		100	
Γ	1.8		2.1		1.5	
p-critical	0.055		0.051		0.063	
A2. College attendance	-0.202	***	-0.138	**	-0.130	
	(0.049)		(0.049)		(0.076)	
Matched pair	218		224		100	
Γ	1.9		1.4		1.2	
p-critical	0.074		0.063		0.056	
B. Labor Market Outcomes						
B1. Employment status	-0.103	**	-0.136	**	0.039	[b,c]
	(0.036)		(0.050)		(0.062)	
Matched pair	448		317		179	
Γ	1.3		1.4		1.0	
p-critical	0.074		0.052		0.229	
B2. Work-related activities	-0.094	**	-0.079		0.106	[b,c]
	(0.037)		(0.051)		(0.062)	
Matched pair	447		317		179	
Γ	1.3		1.1		1.2	
p-critical	0.104		0.062		0.103	
B3. Full-time employment	-0.045		-0.050		-0.054	
	(0.046)		(0.055)		(0.057)	
Matched pair	264		159		111	
Γ	1.0		1.0		1.0	
p-critical	0.113		0.138		0.152	
B4. Weekly wages (logged)	0.013		0.060		0.107	
	(0.065)		(0.081)		(0.077)	
Matched pair	188		117		85	
Γ	1.0		1.0		1.0	
p-critical	0.527		0.419		0.087	
C. Public Assistance Receipt						
C1. Currently on welfare	-0.032		-0.013		0.010	
	(0.040)		(0.065)		(0.068)	
Matched pair	340		150		104	
Γ	1.0		1.0		1.0	
p-critical	0.153		0.450		0.426	
C2. Ever on welfare	-0.047		0.007		-0.010	
	(0.048)		(0.068)		(0.077)	
Matched pair	335		145		102	
Γ	1.0		1.0		1.0	
p-critical	0.103		0.398		0.438	

Notes: Asians are excluded due to their small sample size; For public assistance receipt outcomes, teen mothers are treated as the control group because there are more teen mothers in the sample; For matching estimates for currently on welfare among Hispanics, a caliper of 0.02 instead of 0.01 is used to obtain the Rosenbaum bounds statistics; Bootstrap standard errors in parentheses (300 replications).

[a] Whites differ from Blacks at the .05 level.

[b] Whites differ from Hispanics at the .05 level.

[c] Blacks differ from Hispanics at the .05 level.

*** p<0.001, ** p<0.01, and * p<0.05 (two-tailed tests).

Figure 4.1. Predicted Probability of Teenage Childbearing: Full Sample

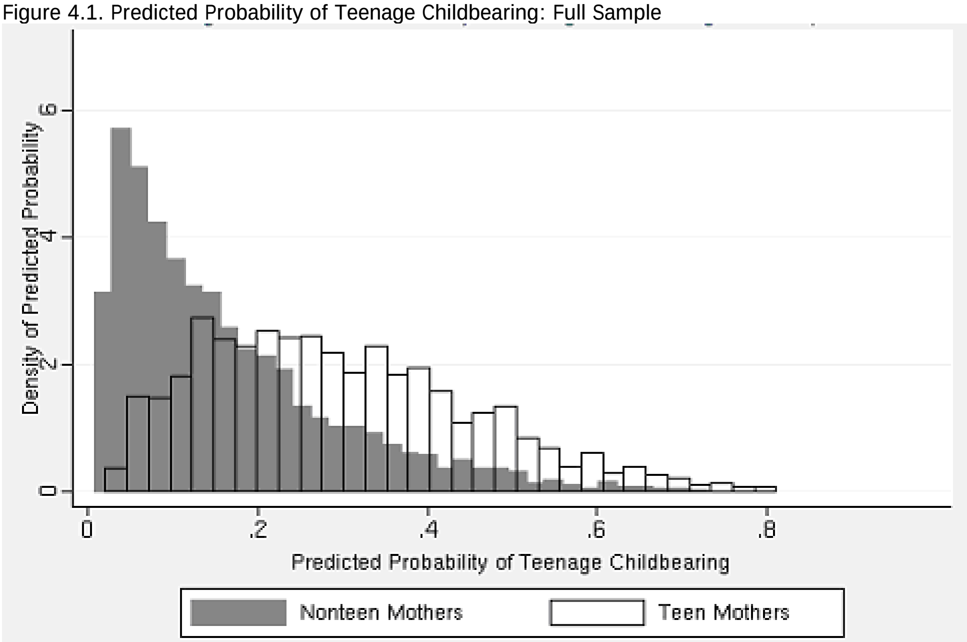
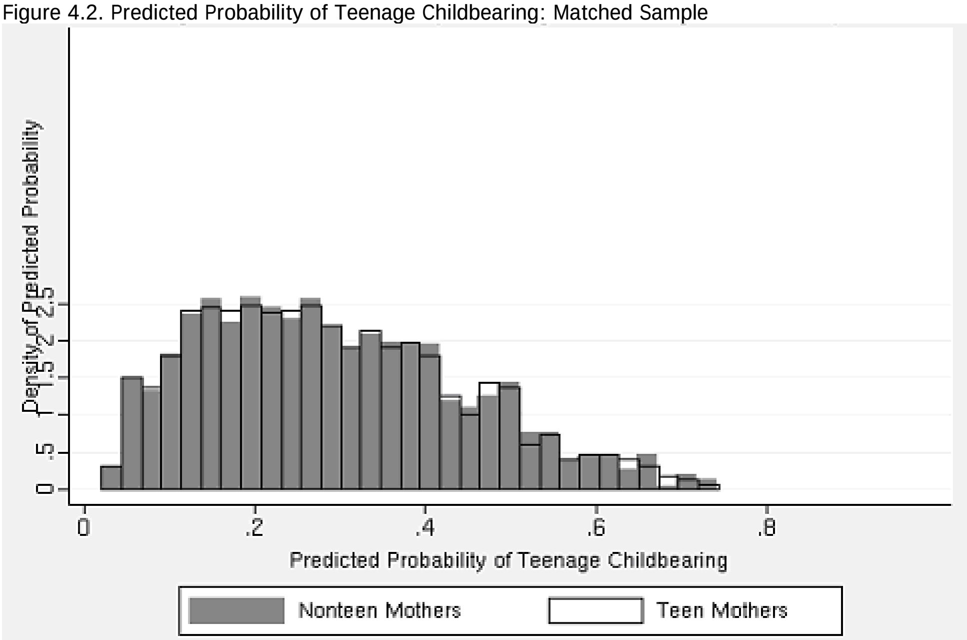


Figure 4.2. Predicted Probability of Teenage Childbearing: Matched Sample



CHAPTER 5

Conclusion

This dissertation has examined the effects on various forms of socioeconomic outcomes of personal traits and habits that have been failed to be fully acknowledged in the scholarship on socioeconomic inequality. Although research has been keen on identifying early predictors of socioeconomic attainment, a systematic view of the linkages between individuals' own attributes formed in childhood and adolescence and subsequent outcomes has yet to come: the role of *individual agency*, which is shaped by social structure and, in turn, redefines it, has been underemphasized. To fill this void in the literature, my research has investigated the role of cognitive and noncognitive skills in multiple life outcomes. Correlated but distinct from cognitive skills, noncognitive skills are conceptualized as enduring dispositions that represent a form of cultural capital. Because both types of skills are highly dependent on socioeconomic background, I hypothesized that they function as a key channel through which intergenerational mobility is associated with socioeconomic inequality. In this chapter, I summarize the major findings of this dissertation, discuss their implications for the literature on socioeconomic inequality, and suggest future directions based on some limitations of this project.

REVIEW OF THE FINDINGS

While the economic return to education has been proposed as a parsimonious explanation for rising wage inequality and its currently high level, this account focuses mainly on between-education group wage inequality and the demand for cognitive skills. To better understand sources of overall wage inequality in light of the increasing economic return to education, Chapter 2 of my dissertation analyzed the impact of cognitive and noncognitive skills on wage inequality by examining the role of skill differences in between- and within-education group wage inequality. There are four findings that emerged from the quantile regression analysis. First, the college wage premium is highly significant, substantial in magnitude, and larger above the median of the wage distribution, even after controlling for cognitive and noncognitive skills as an early predictor of earnings as well as an extensive set of covariates. This finding suggests that the economic return to education is a robust explanation for changes in and levels of wage inequality. The effects of cognitive and noncognitive skills learned in schools, school quality, and educational credentials are reflected in the college wage premium. Second, however, this study revealed that cognitive and noncognitive skills as pre-schooling and pre-labor market factors contribute significantly to reducing both wage differentials between high school and college graduates and wage inequality within college graduates. This finding advances our understanding of how the economic return to education involves within-education group wage inequality, given that the focus on the college wage premium is primarily concerned with between-education group wage inequality. The wage effect of skill differences is mostly driven by the stronger impact of both skills at the upper portion of the wage distribution, which decreases the larger effect of the college premium above the median. In this respect, skill differences provide an important clue to disentangle the polarization of the recent U.S. wage structure, because both cognitive and noncognitive skills are more rewarding at high-paying jobs.

Third, I found that cognitive and noncognitive skills play quite different roles in between- and within-education group wage inequality. The results from the quantile regression interaction models showed that when interacting with the college premium, cognitive skills are more influential in reducing wage differentials between high school and college graduates, whereas noncognitive skills are more pronounced in accounting for wage dispersions within college graduates, especially at the upper tail of the wage distribution. These results suggest that at high-paying jobs, employers who are constantly faced with the principal-agent problem tend to reward a wage premium to college-educated workers if they have high noncognitive skills. Because these jobs allow for more autonomy, college-educated workers with a higher level of cultural capital (e.g., internalized locus of control and self-esteem) are more likely to respond positively to the employer-setting incentive structure. Lastly, cognitive and noncognitive skills as an early predictor of earnings have the long-term effect on between- and within-education group wage inequality. Their effect becomes more salient on both high school-college wage differentials—especially at the upper tail of the wage distribution—and wage inequality within college graduates, as workers approach their prime ages in the labor market. It suggests a cumulative nature of the association between the wage effect of early skill differences and workers' time spent in the labor market.

Chapter 3 of my dissertation explored the extent to which individual patterns of educational assortative mating vary by cognitive and noncognitive skills in order to clarify the relationship between family background and mate sorting by education. Although education has been considered a critical link between mate selection process and socioeconomic inequality, it is difficult to differentiate the role of education as a reflection of family background and its role in developing one's preference for mate choice and serving as marriage markets. The results from multinomial logistic regression models showed that differences in cognitive and noncognitive skills are positively associated with mate selection by education in both adolescence and adulthood. Adolescents with a higher level of skills are more likely to have romantic relationships with the college-bound, although the strength of this association is modest. Adults with a higher level of skills are more likely to transition into marrying the college-educated, and this association is stronger than that found for adolescents. The analysis also found several moderating and/or mediating factors operate in the relationship between skill differences and patterns of educational assortative mating. I identified as such factors their own academic characteristics, parent-child relationship, and school characteristics among adolescents and own economic potential, current living arrangement, and local marriage market characteristics among adults. For adolescents, parent-child relationship plays a salient role in that it tends to discourage them to have romantic relationships regardless of dating partners' level of academic achievement. For adults, their economic potential in particular functions as a mediator of the effect of skill differences, although it does not account fully for their effect.

In addition, this study found a clear gender difference in the role of skill differences: for women, noncognitive skills are more likely to increase the probability of mating the well-educated, whereas for men, it is cognitive skills that primarily do so. This may indicate a normative attitude toward mate choice that regards "smartness" as more attractive to women than to men. Despite this gender difference, these findings suggest that the intergenerational transmission of familial resources affects children's mate selection at its intimate level by not simply investing their educational attainment but also strengthening their cognitive and noncognitive skills.

In Chapter 4, I attempted to make a rigorous causal inference about the effects of teenage childbearing on socioeconomic outcomes, with an emphasis on the role of cognitive and noncognitive skills in teen motherhood and an application of a counterfactual approach. Despite a 30-year debate about the socioeconomic effects of teenage childbearing, finding its "true" effect still has been elusive. This concern stems from 1) theoretical considerations of early motherhood as a harmful event and/or its higher likelihood among disadvantaged young women and 2) methodological challenges against selection bias. Alternative models have been developed, but tend to rely on strong assumptions and unrepresentative samples. This paper extended the extant literature by taking a counterfactual approach using propensity score matching, conducting a sensitivity analysis employing the Rosenbaum bounds to address selection bias on unobserved covariates, and using data from Add Health.

I reported three major findings from this study. First, the propensity score matching results showed that while teen mothers' preexisting disadvantages play a nontrivial role, teen motherhood has significantly negative effects on educational attainment and labor market outcomes with the exception of weekly wages and public assistance receipt. Second, because the matching estimates are grounded on the assumption that selection into teen motherhood on unobserved covariates is ignorable, I also conducted the Rosenbaum bounds method to assess the robustness of these estimates when relaxing the ignorability assumption. The sensitivity analysis suggested that selection bias due to unobserved covariates would have to be very powerful to nullify these findings. Third, the analysis of racial/ethnic differences found that teenage childbearing are detrimental to non-Hispanic whites and to a lesser degree, non-Hispanic blacks, but its effects might be short-lived for Hispanics. While these findings seem in favor of the revisionist view, I cautioned against this interpretation as the negative effect of early motherhood on educational attainment is significant across all racial/ethnic groups. Because a large number of respondents were still enrolled in post-secondary schools at the latest wave of Add Health, the socioeconomic effects of teenage childbearing might be restored if these respondents entered the labor market upon graduation.

IMPLICATIONS OF THE RESEARCH

My research has substantive and methodological implications for the literature on the (re)production of socioeconomic inequality. First of all, with respect to the role of individual skill differences, this study has provided empirical evidence that skills are not unidimensional. Alongside cognitive skills, noncognitive skills make distinctive contribution to explaining between- and within-education group wage inequality, identifying sources of individual patterns of educational assortative mating, and reducing the omitted variables problem contagious in estimations of the effect of teenage childbearing. These findings could enrich the existing literature on socioeconomic inequality by redirecting its attention from cognitive skills to a more diverse set of skills that play a role in social stratification processes. In addition, given that cognitive and noncognitive skills are unevenly distributed among children with different levels of parental socioeconomic status and parenting practice, this study gives credence to research that emphasizes the family as an important institutional actor responsible for socioeconomic inequality. My analysis suggested that intergenerational mobility is connected with socioeconomic inequality in a more complex form, as it relates to investing in children's human capital through their cognitive and noncognitive skill development as well as their educational attainment.

Second, each chapter of this dissertation paves a way to better specify the causes and consequences of socioeconomic inequality. Chapter 2 showed that between- and within-education group wage inequality should be simultaneously taken into consideration in order to advance our understanding of overall wage inequality. By doing so, this study found that the wage effect of cognitive skills runs more noticeably between high school and college graduates and that of noncognitive skills does so within college graduates. In Chapter 3, I investigated the linkages between skill differences and men's and women's education-based mate sorting in both adolescence and adulthood. On the one hand, a closer examination of these linkages among adolescents revealed that although not that strong, social closure persists at the early life stage as far as partnership selection by academic achievement is concerned. On the other hand, the finding of the gender difference in the role of skills implied that while there has been a convergence in the level of educational attainment between partners, one should not ignore the influence of social norms about gender on patterns of educational assortative mating. Chapter 4 drew important public policy implications from the counterfactual analysis of the socioeconomic effect of teenage childbearing. It was found that adolescents' cognitive and noncognitive skills are a significant predictor of the incidence of teenage childbearing. Because these skills are shaped during early childhood, policy interventions need to be made at the early life stages of childhood and pre-adolescence. Furthermore, given that educational attainment is a major domain of socioeconomic disadvantages facing teen mothers, policies aimed at raising the high school completion rate and machining college education more affordable would provide opportunities for teen mothers as well as disadvantaged adolescents as a whole to improve their socioeconomic viability.

Finally, and methodologically, this project has demonstrated how the attempts to clarify the linkages between skill differences and socioeconomic outcomes can benefit from taking careful analytic strategies. For example, as shown in Chapter 2 of my dissertation, quantile regression models offers an alternative approach to traditional regression methods to examine overall wage inequality, where the determinants of wages are theoretically thought to have differential effects across the wage distribution. Also, a propensity matching method in conjunction with the Rosenbaum bounds employed in Chapter 4 provides a unique angle from which to draw a causal inference about the effect of teenage childbearing on socioeconomic consequences, by taking into account selection bias on both observed and unobserved covariates.

LIMITATIONS OF THE RESEARCH AND FUTURE DIRECTIONS

While this dissertation opens a new venue for the research on social stratification with the emphasis on the micro basis of socioeconomic inequality, there are several limitations that warrant further research. First, I treated the role of the family in skill formation as given, based on a vast body of literature that has documented skill differences by family socioeconomic status. However, we do not know much about the mechanisms by which family background affects skill development. A better understanding of the parent-child relationship and the gene-environment interactions is a high priority in this regard. In addition, developing more reliable measures of cognitive and noncognitive skills should be given another priority. Even if I attempted to construct the pre-schooling, pre-labor market measures of these skills to alleviate the measurement error problem, they are limited in scope and depth because the data used in this study provide only verbal and mathematical ability and personality traits. Utilizing multiple indices of aptitude and behavioral traits (e.g., child problem behaviors) seems promising to specify the multidimensionality of skills.

A second limitation is that this project was not able to fully examine how skill differences are associated with various forms of socioeconomic inequality by gender and race/ethnicity. Supplemental analysis attached to Chapter 2 suggested that the wage effect of skill differences is stronger for men than women and to a lesser degree, for whites than minority groups. Future research needs to address whether and how potential institutional arrangements (e.g., female wage penalty or statistical discrimination by race/ethnicity) may suppress the role of skill differences in wage inequality. Chapter 3 indeed examined the gender difference in the role of skill differences in individual patterns of educational assortative mating, but it fell short of dissecting racial/ethnic differences due to data limitations. Such analysis is needed to provide a comprehensive picture of the association between skill levels and education-based mate sorting.

Third, each chapter identifies future research foci that deserve attention. With respect to wage inequality, one could investigate how workers' early skill differences affect their labor market performance under institutional constraints. There may be the linkages between early skill differences and more proximate indicators of wage determination that are heavily influenced by structure of work and occupation. Among such indicators are job-specific skills required at the firm and occupation levels and job mobility. Also, other sources than early skill differences of within-education group wage inequality should be addressed. Given that cognitive and noncognitive skills explain a significant, but modest portion of it, we need to examine other factors such as overeducation, school quality, and the different field of study. Meanwhile, the present study of educational assortative mating poses the "two-sex" problem intrinsic to the mating process, which involves both the "readiness" of individuals with higher levels of skills for mating the well-educated and their “attractiveness” to potential well-educated mate candidates. While I assumed that the process of education-based mate selection goes both directions, this line of research would be inconclusive until new models that take the characteristics of both sexes into account are developed. Finally, the counterfactual analysis of the socioeconomic effect of teenage childbearing presented in Chapter 4 should be understood as an attempt to make the causal inference about its effect more constructive, rather than as the solution to resolve all complex issues regarding selection bias. Further research is needed to clarify the process by which teen motherhood affects subsequent life outcomes, address the heterogeneity among teen mothers, and put the effect of teenage childbearing in the context of macro-level social changes.

In conclusion, this dissertation project joins an emerging literature that underscores the relationship between intergenerational mobility and socioeconomic inequality. Despite an intuitive connection between the two, studies of status attainment and social inequality have been strikingly detached to each other. My research is a sociological effort to bridge them together in order to advance our understanding of the (re)production of inequality. I expect that future research on linking intergenerational mobility and socioeconomic inequality will benefit from incorporating the early skill formation process and its lasting impact on adult life outcomes.

APPENDIX A

The Rosenbaum Bounds Method

The Rosenbaum bounds method is complementary to the estimation of treatment effects using data on matched pairs (Rosenbaum 2002). Although Rosenbaum developed the theory for a more general case, I limit the focus to his treatment of the case of matched pairs (see Chapter 4 of Rosenbaum (2002) and DiPrete and Gangl (2004) for more details).

Test statistics in the family of sign score statistics have the form

$$T = t(Z, r) = \sum_{s=1}^{S} d_s \sum_{i=1}^{2} c_{si} Z_{si,} \tag{A.1}$$

where Z is the variable that designates which of each of the s pairs was treated, and r measures the outcome for each case in the S pairs. Z_{si} equals 1 if a case is treated, and 0 otherwise; c is defined as follows:

$$\begin{aligned}
c_{s1} &= 1, c_{s2} = 0 \quad \text{if} \quad r_{s1} > r_{s2}, \\
c_{s1} &= 0, c_{s2} = 1 \quad \text{if} \quad r_{s1} < r_{s2}, \\
c_{s1} &= 0, c_{s2} = 0 \quad \text{if} \quad r_{s1} = r_{s2}.
\end{aligned}$$

Finally, d_s is the rank of $|r_{s1} - r_{s2}|$ with average ranks used for ties. The product of the c and Z variables causes pairs to be selected where the outcome for the treatment was greater than the outcome for the control. The ranks of these cases are summed and compared with the distribution of the test statistic under the null hypothesis that the treatment has no effect.

In the case where the assignment to treatment is not random, the above test statistic can be bounded. It is assumed that there is an unmeasured variable, U, that affects the probability of receiving the treatment. If we let π_i be the probability that the ith unit receives the treatment, and X is the vector of observed covariates that determine treatment and that also determine the outcome variable, then the following treatment assignment equation applies:

$$\log\left(\frac{\pi_i}{1-\pi_i}\right) = \kappa(X_i) + \gamma U_i, \quad \text{where} \quad 0 \le U_i \le 1. \tag{A.2}$$

Rosenbaum (2002) shows that this relationship implies the following bounds on the ratio of the odds that either of two cases which are matched on *X*—or alternatively on the propensity score *p(X)*—will receive the treatment

$$\frac{1}{\Gamma} \leq \frac{\pi_{s,1}(1-\pi_{s,2})}{\pi_{s,2}(1-\pi_{s,1})} \leq \Gamma, \tag{A.3}$$

where s indexes the matched pair, $s = 1,\ldots,S$, and $\Gamma = \exp(\gamma)$.

Under the assumption that a confounding variable U exists, equation (A.1) becomes the sum of S independent random variables where the sth pair equals d_s with probability

$$p_s = \frac{c_{s1}\exp(\gamma u_{s1}) + c_{s2}\exp(\gamma u_{s2})}{\exp(\gamma u_{s1}) + \exp(\gamma u_{s2})}$$

and equals 0 with probability $1 - p_s$. Define

$$p_s^+ = \begin{cases} 0 & \text{if} \quad c_{s1} = c_{s2} = 0, \\ 1 & \text{if} \quad c_{s1} = c_{s2} = 1, \\ \frac{\Gamma}{1+\Gamma} & \text{if} \quad c_{s1} \neq c_{s2}, \end{cases} \quad \text{and} \quad p_s^- = \begin{cases} 0 & \text{if} \quad c_{s1} = c_{s2} = 0, \\ 1 & \text{if} \quad c_{s1} = c_{s2} = 1, \\ \frac{1}{1+\Gamma} & \text{if} \quad c_{s1} \neq c_{s2}. \end{cases}$$

Rosenbaum (2002) shows that for any specific γ, the null distribution of $t(Z,r)$ is bounded by two known distributions for T^+ and T^- that are attained at values of U, which perfectly predict the signs of c_{si} in equation (A.1), where

$$E(T^+) = \sum_{s=1}^{S} d_s p_s^+ \quad \text{and} \quad \operatorname{var}(T^+) = \sum_{s=1}^{S} d_s^2 p_s^+(1-p_s^+),$$
$$E(T^-) = \sum_{s=1}^{S} d_s p_s^- \quad \text{and} \quad \operatorname{var}(T^-) = \sum_{s=1}^{S} d_s^2 p_s^-(1-p_s^-).$$

One can use these formulae to compute the significance level of the null hypothesis of no treatment effect. For any specific Γ, we compute

$$(T - E(T^+))/\sqrt{\operatorname{var}(T^+)} \quad \text{and} \quad (T - E(T^-))/\sqrt{\operatorname{var}(T^-)}$$

where T is the Wilcoxon signed rank statistic. These two values give bounds of the significance level of a one-sided test for no treatment effect.

APPENDIX B

Covariate Balance Check

Table B1. Covariate Balance Check and Absolute Bias Reduction: Matched Sample

Variable	Non-teen Mothers Mean	Teen Mothers Mean	% Reduction in Absolute Bias
Demographic and family characteristics			
Age at Wave I	15.375	15.410	49.0
White (reference)	0.445	0.427	86.5
Black	0.364	0.354	93.6
Hispanic	0.158	0.182	34.3
Asian	0.033	0.037	92.0
Immigrant generation:			
1st	0.039	0.036	92.7
2nd	0.121	0.124	86.6
3rd+ (reference)	0.840	0.840	100.0
Two-biological parent family (reference)	0.383	0.373	95.5
Step family	0.223	0.230	90.9
Mother only family	0.308	0.309	99.3
Father only family	0.018	0.021	-22.3
Other family structure	0.068	0.067	97.5
Parental education:			
Less than high school	0.191	0.202	88.8
High school graduate	0.379	0.365	85.8
Some college	0.218	0.212	-7710.3
College graduate (reference)	0.176	0.180	98.0
Number of siblings	1.570	1.525	41.0
Individual characteristics			
Parental monitoring	5.097	5.095	97.8
Add Health PVT score	95.117	94.884	96.3
Grade Point Average (GPA)	2.533	2.554	94.7
Rotter/Rosenberg scale	2.170	2.156	89.3
Religiosity	1.952	1.941	85.0
Regular smoking	0.300	0.295	96.1
Frequency of drinking	1.421	1.422	99.8
School characteristics			
Percent white	51.661	50.241	76.2
School type: public	0.973	0.969	92.7
Percent expectations:			
Going to college	77.601	77.617	98.9
Middle class income	62.045	62.152	91.0
West	0.200	0.214	66.6
Midwest	0.230	0.237	-173.4
South	0.490	0.474	85.7
Northeast (reference)	0.079	0.075	94.2
Neighborhood characteristics			
Urbanity	0.249	0.255	81.6
Percent idle	6.093	6.323	82.1
Total unemployment rate	9.159	9.033	92.7
N	1259	1259	

REFERENCES

Allison, P. 1995. *Survival Analysis Using the SAS® System: A Practical Guide*. Cary, NC: SAS Institute Inc.

An, C, R. Haveman, and B. Wolfe. 1993. "Teen Out-of-Wedlock Births and Welfare Receipt: The Role of Childhood Events and Economic Circumstances." *Review of Economics and Statistics* 75(2):195-208.

Angrist, J. and A. Krueger. 1991. "Does Compulsory School Attendance Affect Schooling and Earnings?" *Quarterly Journal of Economics* 106(4):979-1014.

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

Ashenfelter, O. and C. Rouse. 1998. "Income, Schooling, and Ability: Evidence from a New Sample of Identical Twins." *Quarterly Journal of Economics* 113(1):253-284.

Autor, D., L. Katz, and M. Kearney. 2005. "Rising Wage Inequality: The Role of Composition and Prices." NBER Working Paper No. 11628.

Autor, D., L. Katz, and M. Kearney. 2006. "The Polarization of the U.S. Labor Market." *American Economic Review Papers and Proceedings* 96(2):189-194.

Becker, G. 1981. *A Treatise on the Family*. Cambridge, MA: Harvard University Press.

______. 1993. *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*. Chicago, IL: University of Chicago Press.

Bellah, R. et al. 1985. *Habit of the Heart: Individualism and Commitment in American Life*. Berkeley, CA: University of California Press.

Belzil, C. and J. Hansen. 2002. "Unobserved Ability and the Return to Schooling." *Econometrica* 70(5):2075-2091.

Bernhardt, A. et al. 2001. *Divergent Paths: Economic Mobility in the New American Labor Market*. New York, NY: Russell Sage Foundation.

Blackwell, D. 1998. "Marital Homogamy in the United States: The Influence of Individual and Paternal Education." *Social Science Research* 27(2):159-188.

Blau, P. and O. Duncan. 1967. *The American Occupational Structure*. New York, NY: Free Press.

Bourdieu, P. 1977. *Outline of a Theory of Practice*. Cambridge, UK: Cambridge University Press.

______. 1984. *Distinction: a Social Critique of the Judgment of Taste*. Cambridge, MA: Harvard University Press.

______. 2002. "Pierre Bourdieu on Marriage Strategies." *Population and Development Review* 28(3):549-558.

Bourdieu, P. and L. Wacquant. 1992. *An Invitation to Reflexive Sociology*. Chicago, IL: University of Chicago Press.

Bowles, S. and H. Gintis. 2000. "Does Schooling Raise Earnings by Making People Smarter?" In *Meritocracy and Economic Inequality*, eds. Arrow, K., Bowles, S., and S. Durlauf, pp. 118-136. Princeton, NJ: Princeton University Press.

Bowles, S. and H. Gintis. 2002a. "Schooling in Capitalist America Revisted." *Sociology of Education* 75(1):1-18.

Bowles, S. and H. Gintis. 2002b. "The Inheritance of Inequality." *Journal of Economic Perspectives* 16(3):3-30.

Bowles, S., Gintis, H., and M. Osborne. 2001. "The Determinants of Earnings: A Behavioral Approach." *Journal of Economic Literature* 39(4):1137-1176.

Breier, R. 1995. "Social Structure and the Phenomenology of Attainment." *Annual Review of Sociology* 21:115-136.

Buchinsky, M. 1994. "Changes in the U.S. Wage Structure 1963-1987: Application of Quantile Regression." *Econometrica* 62(2):405-458.

______. 1998. "Recent Advances in Quantile Regression Models: A Practical Guideline for Empirical Research." *Journal of Human Resources* 33(1):88-126.

Bumpass, L., J. Sweet, and A. Cherlin. 1991. "The Role of Cohabitation in Declining Rates of Marriage." *Journal of Marriage and the Family* 53(4):913-927.

Cameron, S, and J. Heckman. 1993. "The Nonequivalence of High School Equivalents." *Journal of Labor Economics* 11(1):1-47.

Cao, J., E. Stromsdorfer, and G. Weeks. 1996. "The Human Capital Effect of General Education Development Certificates on Low Income Women." *The Journal of Human Resources* 31(1):206-228.

Card, D. 1995a. "Using Geographic Variation in College Proximity to Estimate the Return to Schooling." In *Aspects of Labour Market Behavior: Essays in Honour of John*

Vanderkamp, eds. Christofides, L., E.K. Grant, and R. Swidinsky, pp. 201-222. Toronto, Canada: University of Toronto Press.

______. 1995b. “Earnings, Schooling, and Ability Revisited.” In *Research in Labor Economics*, vol. 14, ed. Polachek, S. Greenwich, CT: JAI Press.

______. 1999. “The Causal Effect of Education on Earnings.” In *Handbook of Labor Economics*, vol. 3, eds. Ashenfelter, O. and D. Card. Amsterdam, Netherland: Elsevier.

Carneiro, P. and J. Heckman 2003. “Human Capital Policy.” In *Inequality in America: What Role for Human Capital Policy?*, eds. Heckman, J. and A. Krueger. Cambridge, MA: MIT Press.

Carver, Karen, Kara Joyner, and J. Richard Udry. 2003. “National Estimates of Adolescent Romantic Relationships.” In *Adolescent Romantic Relations and Sexual Behavior: Theory, Research, and Practical Implications*, ed. Florsheim. P., pp. 23-56. Mahwah, NJ: Lawrence Erlbaum and Associates.

Cawley et al. 2000. “Understanding the Role of Cognitive Ability in Accounting for the Recent Rise in the Return to Education.” In *Meritocracy and Economic Inequality*, eds. Arrow, K., Bowles, S., and S. Durlauf, pp. 230-266. Princeton, NJ: Princeton University Press.

Center for Human Resource Research. 2002. *A Guide to the 1979-2002 National Longitudinal Survey of Youth Data*. Columbus, OH: The Ohio State University.

Chantala, K. 2001. “Introduction to Analyzing Add Health Data.” Carolina Population Center, University of North Carolina at Chapel Hill, Chapel Hill, NC.

Cheng, S. and S. Long. 2007. “Testing for IIA in the Multinomial Logit Model.” *Sociological Methods and Research* 35(4):583-600.

Cherlin, A. 2001. “New Developments in the Study of Nonmarital Childbearing.” In *Out of Wedlock: Causes and Consequences of Nonmarital Fertility*, ed. Wu, L. and B. Wolfe, pp. 390-402. New York, NY: Russell Sage Foundation.

______. 2004. “The Deinstitutionalization of American Marriage.” *Journal of Marriage and Family* 66(4):848-861.

Chevalier, A. and T. Viitanen. 2003. “The Long-Run Labour Market Consequences of Teenage Motherhood in Britain.” *Journal of Population Economics* 16(2):323-343.

Cohen, J. 1977. “Sources of Peer Group Homogeneity.” *Sociology of Education* 50(4):227-241.

Coleman, J. 1990. *Foundations of Social Theory*. Cambridge, MA: Harvard University Press.

Condron, D. 2007. “Stratification and Educational Sorting: Explaining Ascriptive Inequalities in Early Childhood Reading Group Placement.” *Social Problems* 54(1):139-160.

Corcoran, M. 1995. “Rags to Rags: Poverty and Mobility in the United States.” *Annual Review of Sociology* 21:237-267.

Dehejia, R. and S. Wahba. 2002. “Propensity Score Matching Methods for Non-Experimental Causal Studies.” *Review of Economics and Statistics* 84(1):151-161.

Diamond, A. and J. Sekhon. 2005. “Genetic Matching for Estimating Causal Effects: A General Multivariate Matching Method for Achieving Balance in Observational Studies.” Working Paper, Travers Department of Political Science, UC Berkeley.

DiMaggio, P. 1982. “Cultural Capital and School Success: the Impact of Status-Culture Participation on the Grades of U.S. High-School Students.” *American Sociological Review* 47(2):189-201.

DiMaggio, P. and J. Mohr. 1985. “Cultural Capital, Educational Attainment, and Marital Selection.” *American Journal of Sociology* 90(6):1231-1261.

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

Duckworth, A. and M. Seligman. 2005. “Self-Discipline Outdoes IQ in Predicting Academic Performance of Adolescents.” *Psychological Science* 16(12):939-944.

Dunifon, R. and G. Duncan. 1998. “Long-Run Effects of Motivation on Labor Market Success.” *Social Psychology Quarterly* 61(1):33-48.

Dunifon, R., G. Duncan, and J. Brooks-Gunn. 2001. “As Ye Sweep, So Shall Ye Reap.” *American Economic Review Papers and Proceedings* 91(2):150-154.

Farkas, G. 2003. “Cognitive Skills and Noncognitive Traits and Behaviors in Stratification Processes.” *Annual Review of Sociology* 29:541-562.

Farkas, G., R. Grobe, D. Sheehan, and Y. Shuan. 1990. “Cultural Resources and School Success: Gender, Ethnicity, and Poverty Groups within an Urban School District.” *American Sociological Review* 55(1):127-142.

Fernández, R., N. Guner, and J. Knowles 2005. “Love and Money: A Theoretical and Empirical Analysis of Household Sorting and Inequality.” *Quarterly Journal of Economics* 120(1):273-344.

Furstenberg, F. 1991. “As the Pendulum Swings: Teenage Childbearing and Social Concern.” *Family Relations* 40(2):127-138.

Geronimus, A. 1991. “Teenage Childbearing and Social and Reproductive Disadvantage: The Evolution of Complex Questions and the Demise of Simple Answers.” *Family Relations* 40(4):463-471.

Geronimus, A. and S. Korenman. 1992. “The Socioeconomic Consequences of Teenage Childbearing Reconsidered.” *Quarterly Journal of Economics* 107(4):1187-1214.

Geronimus, A. and S. Korenman. 1993. “The Socioeconomic Costs of Teenage Childbearing: Evidence and Interpretation.” *Demography* 30(2):281-290.

Geronimus, A., S. Korenman, and M. Hillemeier. 1994. “Does Young Maternal Age Adversely Affect Child Development? Evidence from Cousin Comparisons in the United States.” *Population and Development Review* 29(3):585-609.

Goldscheider, F. and L. Waite. 1986. “Sex Differences in the Entry into Marriage.” *American Journal of Sociology* 92(1):91-109.

Goldstein, J. and C. Kenney. 2001. “Marriage Delayed or Marriage Forgone? New Cohort Forecasts of First Marriage for U.S. Women.” *American Sociological Review* 66(4):506-519.

Green, F., S. Machin, and D. Wilkenson. 1998. “The Meaning and Determinants of Skill Shortages.” *Oxford Bulletin of Economics and Statistics* 60(2):165-187.

Grogger, J. and S. Bronars. 1993. “The Socioeconomic Consequences of Teenage Childbearing: Findings from a Natural Experiment.” *Family Planning Perspectives* 25(4):156-161&174.

Hao, L. and D. Naiman. 2007. *Quantile Regression*. New York, NY: Sage Publications.

Harding, D. 2003. “Counterfactual Models of Neighborhood Effects: The Effect of Neighborhood Poverty on Dropping Out and Teenage Pregnancy.” *American Journal of Sociology* 109(3):676-719.

Harris, K. M. 1999. “The Health Status and Risk Behavior of Adolescents in Immigrant Families.” In *Children of Immigrants: Health, Adjustment, and Public Assistance*, ed. Hernandez, D., pp. 286-347. Washington, D.C.: National Academy Press.

Harris, K. M. et al. 2003. “The National Longitudinal Study of Adolescent Health: Research Design.” [WWW document]. URL: www.cpc.unc.edu/projects/addhealth/design.

Heckman, J., H. Ichimura, and P. Todd. 1998. "Matching as an Econometric Evaluation Estimator." *Review of Economic Studies* 65:261-94.

Heckman, J. and E. Vytlacil. 1999. "Local Instrumental Variables and Latent Variable Models for Identifying and Bounding Treatment Effects." *Proceedings of the National Academy of Sciences of the United States of America* 96(8):4730-4734.

Heckman, J. and Y. Rubinstein. 2001. "The Importance of Noncognitive Skills: Lessons from the GED Testing Program." *American Economic Review Papers and Proceedings* 91(2):145-149.

Heckman, J., J. Stixrud, and S. Urzua. 2006. "The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior." *Journal of Labor Economics* 24(3):411-482.

Herrnstein, R. and C. Murray. 1994. *The Bell Curve: Intelligence and Class Structure in American Life*. New York, NY: Free Press.

Hofferth, S. and C. Hayes. 1987. *Risking the Future*. Washington, DC: National Academy Press.

Hoffman, S., M. Foster, and F. Furstenberg. 1993a. "Reevaluating the Costs of Teenage Childbearing." *Demography* 30(1):1-13.

Hoffman, S., M. Foster, and F. Furstenberg. 1993b. "Reevaluating the Costs of Teenage Childbearing: Response to Geronimus and Korenman." *Demography* 30(2):291-296.

Hoffman, S. 1998. "Teenage Childbearing Is Not So Bad After All… Or Is It? A Review of the New Literature." *Family Planning Perspectives* 30(5):236-239&243.

Holland, P. 1986. "Statistics and Causal Reasoning." *Journal of the American Statistics Association* 81:945-960.

Holmlund, Helena. 2005. "Estimating Long-Term Consequences of Teenage Childbearing: An Examination of the Siblings Approach." *Journal of Human Resources* 40(3):716-743.

Holzer, H. 1996. *What Employers Want: Job Prospects for Less-Educated Workers*. New York: Russell Sage Foundation.

Hotz, VJ., S. McElroy, and S. Sanders. 1997. "The Costs and Consequences of Teenage Childbearing for the Mothers and the Government." In *Kids Having Kids*, ed. Maynard, R., pp. 55-94. Washington, DC: The Urban Institute Press.

Hotz, VJ., S. McElroy, and S. Sanders. 2005. "Teenage Childbearing and Its Life Cycle Consequences: Exploiting a Natural Experiment." *Journal of Human Resources* 40(3):683-715.

Hout, M. 2004. "Social Mobility and Inequality: A Review and an Agenda." In *Social Inequality*, ed. Neckerman, K., pp. 969-987. New York, NY: Russell Sage Foundation.

Imai, K. and D. van Dyk. 2004. "Causal Inference with General Treatment Regimes: Generalizing the Propensity Score." *Journal of the American Statistical Association* 99(467):854-866.

Imbens, G. 2000. "The Role of the Propensity Score in Estimating Dose–Response Function." *Biometrika* 87(3):706-710.

Imbens, G. and J. Angrist. 1994. "Identification and Estimation of Local Average Treatment Effects." *Econometrica* 62(2):467-475.

Jencks, C. et al. 1979. *Who Gets Ahead?* New York, NY: Basic Books.

Joyner, K. and J. R. Udry. 2000. "You Don't Bring Me Anything but Down: Adolescent Romance and Depression." *Journal of Health and Social Behavior* 41(4): 369-391.

Juhn, C., K. Murphy, and B. Pierce. 1993. "Wage Inequality and the Rise in Returns to Skill." *Journal of Political Economy* 101(3):410-442.

Kalmijn, M. 1991a. "Status Homogamy in the United States." *American Journal of Sociology* 97(2):496-523.

______. 1991b. "Shifting Boundaries: Trends in Religious and Educational Homogamy." *American Sociological Review* 56(6):786-800.

______. 1994. "Assortative Mating by Cultural and Economic Occupational Status." *American Journal of Sociology* 100(2):422-452.

______. 1998. "Intermarriage and Homogamy: Causes, Patterns, Trends." *Annual Review of Sociology* 24:395-421.

Kandel, D. 1978. "Homophily, Selection, and Socialization in Adolescent Friendship." *American Journal of Sociology* 84(2):427-436.

Katz, L. and K. Murphy. 1992. "Changes in Relative Wages, 1963-1987: Supply and Demand Factors." *Quarterly Journal of Economics* 107(1):35-78.

Kim, C-H. and A. Sakamoto. 2008. "The Rise of Intra-Occupational Wage Inequality in the United States, 1983 to 2002." *American Sociological Review* 73(1):129-157.

Klepinger, D., S. Lundberg, and R. Plotnick. 1999. “How Does Adolescent Fertility Affect the Human Capital and Wages of Young Women?” *Journal of Human Resources* 34(3):421-448.

Koenker, R. and G. Bassett. 1978. “Regression Quantiles.” *Econometrica* 46(1):33-50.

Koenker, R. and K. Hallock. 2001. “Quantile Regression.” *Journal of Economic Perspectives* 15(4):143-156.

Korenman, S., R. Kaestner, and T. Joyce. 2001. “Unintended Pregnancy and the Consequences of Nonmarital Childbearing.” In *Out of Wedlock: Causes and Consequences of Nonmarital Fertility*, ed. Wu, L. and B. Wolfe, pp. 259-286. New York, NY: Russell Sage Foundation.

Kuhn, P. and C. Weinberger. 2005. “Leadership Skills and Wages.” *Journal of Labor Economics* 23(3):395-436.

Lamont, M. and A. Lareau. 1988. “Cultural Capital: Allusions, Gaps and Glissandos in Recent Theoretical Developments.” *Sociological Theory* 6(2):153-168.

Lareau, A. 2002. “Invisible Inequality: Social Class and Child Rearing in Black Families and White Families.” *American Sociological Review* 67(5):747-776.

Lemieux, T. 2006. “Increasing Residual Wage Inequality: Composition Effects, Noisy Data, or Rising Demand for Skill?” *American Economic Review* 96(3):461-498.

Levine, D. and G. Painter. 2003. “The Schooling Costs of Teenage Out-of-Wedlock Childbearing: Analysis with a Within-School Propensity-Score-Matching Estimator.” *Review of Economics and Statistics* 85(4):884-900.

Lewis, S. and V. Oppenheimer. 2000. “Educational Assortative Mating across Markets: Non-Hispanic Whites in the United States.” *Demography* 37(1):29-40.

Lichter, D. et al. 1992. “Race and the Retreat from Marriage: A Shortage of Marriageable Men?” *American Sociological Review* 57(6):781-799.

Lipset, M. and R. Bendix. 1959. *Social Mobility in Industrial Society*. Berkeley, CA: University of California Press.

Lloyd, K. and S. South. 1996. “Contextual Influences on Young Men’s Transition to First Marriage.” *Social Forces* 74(3):1097-1119.

Mare, R. 1991. “Five Decades of Educational Assortative Mating.” *American Sociological Review* 56(1):15-32.

Martin, J. et al. 2006. "Births: Final Data for 2004." *National Vital Statistics Reports* 55(1). Hyattsville, MD: National Center for Health Statistics.

Martins, P. and P. Pereira. 2004. "Does Education Reduce Wage Inequality? Quantile Regression Evidence from 16 Countries." *Labour Economics* 11(3):355-371.

Massey, D. and N. Denton. 1993. *American Apartheid*. Cambridge, MA: Harvard University Press.

McCall, L. 2000. "Gender and the New Inequality: Explaining the College/Non-College Wage Gap in U.S. Labor Markets." *American Sociological Review* 65(2):234-255.

McPherson, M., L. Smith-Lovin, and J. Cook. 2001. "Birds of a Feather: Homophily in Social Networks." *Annual Review of Sociology* 27:415-444.

Meier, A. 2007. "Adolescent First Sex and Subsequent Mental Health." *American Journal of Sociology* 112(6):1811-1847.

Moffitt, R. 2005. "Remarks on the Analysis of Causal Relationships in Population Research." *Demography* 42(1):91-108.

Morgan, S. 2001. "Counterfactuals, Causal Effect Heterogeneity, and the Catholic School Effect on Learning." *Sociology of Education* 74(3):341-74

Morgan, S. and D. Harding. 2006. "Matching Estimators of Causal Effects: Prospects and Pitfalls in Theory and Practice." *Sociological Methods and Research* 35(1):3-60.

Morgan, S. and C. Winship. 2007. *Counterfactuals and Causal Inference: Methods and Principles for Social Research*. New York, NY: Cambridge University Press.

Morris, M. and B. Western 1999. "Inequality in Earnings at the Close of the Twentieth Century." *Annual Review of Sociology* 25:623-657.

Mouw, T. and A. Kalleberg. 2008. "Occupations and the Structure of Wage Inequality in the United States, 1980s-2000s." Working Paper. University of North Carolina at Chapel Hill.

Murnane, R., Willett, J., and F. Levy. 1995. "The Growing Importance of Cognitive Skills in Wage Determination." *Review of Economics and Statistics* 77(2):251-266.

Murphy, K. and F. Welch. 1992. "The Structure of Wages." *Quarterly Journal of Economics* 107(1):285-326.

National Association of Colleges and Employers. 2000. "Ideal Candidate Has Top-Notch Interpersonal Skills, Say Employers." http://www.naceweb.org

Neckerman, K. and F. Torche. 2007. “Inequality: Causes and Consequences.” *Annual Review of Sociology* 33:335-357.

Olsen, R. and G. Farkas. 1989. “Endogenous Covariates in Duration Models and the Effect of Adolescent Childbirth on Schooling." *Journal of Human Resources* 24(1):39-53.

Oppenheimer, V. 1988. “A Theory of Marriage Timing.” *American Journal of Sociology* 94(3):563-591.

Oppenheimer, V., H. Blossfeld, and A. Wackerow. 1995. “United States of America” In *The New Role of Women: Family Formation in Modern Societies*, ed. Blossfeld, H, pp. 150-173. Boulder, CO: Westview.

Pencavel, J. 1998. “Assortative Mating by Schooling and the Work Behavior of Wives and Husbands.” *American Economic Review* 88(2):326-329.

Plotnick, R. 1992. “The Effects of Attitudes on Teenage Premarital Pregnancy and Its Resolution.” *American Sociological Review* 57(6):800-811.

Qian, Z. and S. Preston. 1993. “Changes in American Marriage, 1972-1987: Availability and Forces of Attraction by Age and Education.” *American Sociological Review* 58(4):482-495.

Raley, R., S. Crissey, and C. Muller. 2007. “Of Sex and Romance: Late Adolescent Relationships and Young Adult Union Formation.” *Journal of Marriage and Family* 69(4):1210-1226.

Ribar, D. 1994. “Teenage Fertility and High School Completion.” *Review of Economics and Statistics* 76(3):413-424.

______. 1999. “The Socioeconomic Consequences of Young Women’s Childbearing: Reconciling Disparate Evidence.” *Journal of Population Economics* 12:547-565.

Rindfuss, R. 1991. “The Young Adult Years: Diversity, Structural Change, and Fertility.” *Demography* 28(4):493-512.

Rosenbaum, J. 2001. *Beyond College for All: Career Paths for the Forgotten Half*. New York, NY: Russell Sage Foundation.

Rosenbaum, P. 2002. *Observational Studies*. New York, NY: Springer.

Rosenbaum, P. and D. Rubin. 1983. “The Central Role of the Propensity Score in Observational Studies for Causal Effects.” *Biometrika* 70:41-55.

Rosenberg, M. 1965. *Society and the Adolescent Self-Image*. Princeton, NJ: Princeton University Press.

Rosenfeld, M. 2005. “A Critique of Exchange Theory in Mate Selection.” *American Journal of Sociology* 110(5):1284-1325.

Rotter, J. 1966. “Generalized Expectancies for Internal Versus External Control of Reinforcement.” In *Applications of a Social Learning Theory of Personality*, ed. Rotter, J., J. Chance, and E.J. Phares, pp. 260-295. New York, NY: Holt, Rinehart and Winston, Inc.

Rubin, D. 1977. “Assignment to a Treatment Group on the Basis of a Covariate.” *Journal of Educational Statistics* 2(1):1-26.

Rytina, S. 1992. “Scaling the Intergenerational Continuity of Occupation: Is Occupational Inheritance Ascriptive After All?” *American Journal of Sociology* 97(6):1658-1688.

Schoen, R. and R. Weinick. 1993. “Partner Choice in Marriages and Cohabitations.” *Journal of Marriage and the Family* 55:408-414.

Schwartz, C. and R. Mare. 2005. “Trends in Educational Assortative Marriage from 1940 to 2003.” *Demography* 42(4):621-646.

Stevens, G. 1991. “Propinquity and Educational Homogamy.” *Sociological Forum* 6(4):715-726.

Sweeney, M. 2002. “Two Decades of Family Change: The Shifting Economic Foundations of Marriage.” *American Sociological Review* 67(1):132-147.

Swidler, A. 1986. “Culture in Action: Symbols and Strategies.” *American Sociological Review* 51(2):273-286.

Taber, C. 2001. “The Rising College Premium in the Eighties: Return to College or Return to Unobserved Ability?” *Review of Economic Studies* 68(3):666-691.

Thornton, A., Axinn, W., and J. Teachman. 1995. “The Influence of School Enrollment and Accumulation on Cohabitation and Marriage in Early Adulthood.” *American Sociological Review* 60(5):762-774.

Udry, R. and K. Chantala. 2003. “Missing School Dropouts in Surveys Does Not Bias Risk Estimates.” *Social Science Research* 32(2):294-311.

Upchurch, D. and J. McCarthy. 1990. “The Timing of a First Birth and High School Completion.” *American Sociological Review* 55(2):224-234.

U.S. Bureau of the Census. 1998. “First Findings from the EQW National Employer Survey.” EQW Catalog RE01.

Ventura, S. et al. 2004. “Estimated Pregnancy Rates for the United States, 1990-2000: An Update.” *National Vital Statistics Reports*. Hyattsville, MD: National Center for Health Statistics.

Wang, H., G. Kao, and K. Joyner. 2006. “Stability of Interracial and Intraracial Romantic Relationships among Adolescents.” *Social Science Research* 35(4):435-453.

Wilson, W. 1987. *The Truly Disadvantaged*. Chicago, IL: University of Chicago Press.

Winship, C. and R. Mare. 1992. “Models for Sample Selection Bias.” *Annual Review of Sociology* 18:327-350.

Winship, C. and S. Morgan. 1999. “The Estimation of Causal Effects from Observational Data.” *Annual Review of Sociology* 25:659-706.

Wilson, W. 1987. *The Truly Disadvantaged*. Chicago, IL: University of Chicago Press.

Wu, L. and B. Martinson. 1993. “Family Structure and the Risk of a Premarital Birth.” *American Sociological Review* 58(2):210-232.

Wu, L. and B. Wolfe. 2001. *Out of Wedlock: Causes and Consequences of Nonmarital Fertility*. New York, NY: Russell Sage Foundation.

Zimmerman, D. 1992. “Regression toward Mediocrity in Economic Stature.” *American Economic Review* 82(3):409-429.