Consumer Risk Preferences and Higher Education Enrollment Decisions

DISSERTATION

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy
in the Graduate School of The Ohio State University

By

Stuart J. Heckman

Graduate Program in Family Resource Management

The Ohio State University

2014

Dissertation Committee:

Catherine P. Montalto, Ph.D., Advisor

Sherman D. Hanna, Ph.D.

Rachel E. Dwyer, Ph.D.

Stephen R. Cosslett, Ph.D.

Copyrighted by
Stuart J. Heckman
2014

Abstract

Concerns that consumers are making poor choices in higher education are widespread. Popular media, administrators, and policymakers are primarily concerned about student loans, defaults, student financial wellness, and labor market outcomes. However, a key aspect that is overlooked in the national conversation regarding consumer higher education choices is that human capital investments are risky. That is, there is a great deal of risk involved in the decision to pursue higher education. Since the field of behavioral economics suggests that, in general, consumers are prone to make mistakes in choices when risk is involved, the lack of research on higher education enrollment as a risky choice is surprising. The extent to which consumers understand the risk involved and make rational choices has not been thoroughly explored in the literature regarding human capital investments.

Therefore, the purpose of this research was to investigate the *ceteris paribus* effect of consumer risk preferences on the decision to enroll in higher education. A sample from the 1997 cohort of the National Longitudinal Study of Youth (NLSY97) was analyzed by using logistic regression to model the likelihood of higher education enrollment among young adults. Using the NLSY97 allowed for strong individual-level controls in the empirical model, including explanatory variables that have been consistently demonstrated in the literature to predict college enrollment. In addition to the

standard individual-level controls, this study advanced the understanding of enrollment decisions by including measures for time preferences, subjective perceptions of risk in pursuing higher education, and risk preferences, all of which were identified as important predictors in a risky human capital theoretical model. Since the literature regarding human capital accumulation and the returns to education is vast and spans multiple disciplines, this research also contributes to the literature by providing a thorough review of research and theoretical models across disciplines.

The results of the logistic regression indicated that consumer risk preferences have a significant effect on the likelihood of enrollment. Specifically, there was a positive relationship between risk tolerance and the likelihood of enrollment. This finding was robust to a number of empirical specifications, including sample restrictions, measurement of risk tolerance, and timing of variable measurement. Other findings were generally consistent with previous literature, including positive and significant effects of academic ability, parental education, and family net worth and income. Women were also significantly more likely to enroll than men and Blacks were significantly more likely to enroll than Whites.

The finding that more risk tolerant individuals are more likely to enroll in higher education implies that to consumers, attending higher education seems riskier than entering the labor market directly. A number of productive areas for future research are identified, including the need for qualitative research to further investigate risk perceptions regarding higher education and the extent to which consumers are making informed choices. A number of implications for higher education stakeholders and policymakers are also discussed. Lastly, this research has highlighted several

methodological issues regarding higher education decision making research using the NLSY97.

To my wife

Acknowledgments

This work would not have been possible without a number of people who have provided tremendous support and guidance throughout my graduate education and in the completion of this dissertation.

I wish to thank my advisor, Dr. Catherine Montalto, for her careful thinking, enthusiasm for rigorous research, and excellent questions that guided me in the development of this project. She is an outstanding researcher and teacher, and I feel privileged to have worked with her at Ohio State.

Throughout this research and my graduate program, I have benefited greatly from Dr. Sherman Hanna's commitment to graduate student education and deeply appreciate his investment in my research. I also wish to thank the rest of my committee for their help throughout this project and their willingness to serve in this capacity outside of their home departments. I especially thank Dr. Rachel Dwyer for her help with the dataset and Dr. Stephen Cosslett for his help with the methodology.

I also wish to thank the faculty members in the Department of Human Sciences at Ohio State, especially, Dr. Robert Scharff and Dr. Kathryn Stafford. Both Dr. Scharff and Dr. Stafford were very helpful in asking questions early in the process to help develop this research. I am also grateful for the graduate support that I have received throughout my graduate education from the university, the Consumer Sciences program, the

Department of Human Sciences, the Student Wellness Center, and the Center for Higher Education Enterprise.

Lastly, and most importantly, I would like to thank my family for their support throughout my graduate education. I especially wish to thank my wife, Heidi, for her patience during my graduate education and while I completed this dissertation. This process has been challenging in a variety of ways and this dissertation would not have been possible without her support.

Vita

2010	B.S. Family Studies and Human Services, Kansas State University
2011 to 2012	Graduate Teaching Associate, Department of Consumer Sciences, The Ohio State University
2012	M.S. Family Resource Management, The Ohio State University
Spring 2013	Graduate Research Associate, Student Wellness Center, The Ohio State University
Fall 2013	Graduate Teaching Associate, Department of Human Sciences, The Ohio State University
2014 – Present	Graduate Research Associate, Center for Higher Education Enterprise, The Ohio State University
2014	Ph.D. Family Resource Management, The Ohio State University
Honors and Awards.....	University Fellowship, The Ohio State University (2010-2011) Graduate Teaching Award, Department of Consumer Sciences, The Ohio State University (2012, 2013)

Graduate Student Excellence, Department of
Human Sciences, The Ohio State University
(2014)

Publications

- Letskiewicz, J.C., Lim H., Heckman, S.J., Bartholomae, S., Fox, J.J., & Montalto, C.P. (in press). The path to graduation: Factors predicting on-time graduation rates. *Journal of College Student Retention: Research, Theory & Practice*.
- Heckman, S.J., Lim, H., & Montalto, C.P. (2013). Factors related to financial stress among college students. *Consumer Interests Annual*, 59.
- Heckman, S.J. (2013). Book review: Tools and techniques of financial planning by Leimberg, Satinsky, Doyle, and Jackson. *Family and Consumer Sciences Research Journal*, 41 (4).
- Heckman, S.J., & Hanna, S.D. (2012). Factors related to saving behavior among low-income households in the 1995-2007 Survey of Consumer Finances. *Proceedings of the Academy of Financial Services*.
- Heckman, S.J. & Grable, J.E. (2011). Testing the role of parental debt attitudes, student income, dependency status, and financial knowledge have shaping financial self-efficacy among college students. *College Student Journal*, 45 (1).
- Heckman, S.J. (2009). Personal finance knowledge and self-efficacy among college students. *Proceedings of the Association for Financial Counseling and Planning Education*.

Fields of Study

Major Field: Family Resource Management

Minor Fields: Economics, Statistical Data Analysis

Table of Contents

Abstract	ii
Dedication	v
Acknowledgments	vi
Vita	viii
List of Tables	xiv
List of Figures	xvi
Chapter 1: Introduction	1
1.1 Background	1
1.2 Purpose	4
1.3 Significance	6
Chapter 2: Literature Review	8
2.1 Risk and Consumer Risk Preferences	9
2.2 Risk and Education Decisions	12
2.3 College-Going Behavior	16
2.4 Persistence	20
2.5 Returns to Education	24

2.6 Summary and Gap	28
Chapter 3: Review of Theoretical Approaches	30
3.1 Economic Approaches	31
3.1.1 Human Capital Theory	31
3.1.2 Signaling Theory	34
3.1.3 Expected Utility Theory	37
3.1.4 Risky Human Capital Models	40
3.1.5 Summary of Economic Approaches	43
3.2 Sociological Approaches	44
3.2.1 Social and Cultural Capital	44
3.2.2 Relative Risk Aversion Theory	46
3.2.3 Summary of Sociological Approaches	50
3.3 Higher Education Approaches	50
3.3.1 Student College Choice Process	50
3.3.2 Integrated Model of Student College Choice	53
3.3.3 Financial Nexus Model	55
3.3.4 Summary of Higher Education Approaches	59
3.4 Summary	59
Chapter 4: Theory	62

4.1 Theoretical Model	62
4.2 Empirical Model.....	64
4.3 Hypotheses	66
Chapter 5: Research Design and Method.....	67
5.1 Data	67
5.2 Sample.....	69
5.3 Measurement of Variables	71
5.3.1 Dependent Variable	71
5.3.2 Independent Variables	72
5.4 Research Method.....	86
5.4.1 Model 1.....	86
5.4.2 Gender Differences.....	88
5.5 Sensitivity Analysis.....	90
5.5.1 Missing Data.....	90
5.5.2 Measurement of Risk Tolerance.....	94
5.5.3 Timing of Variable Measurement.....	96
Chapter 6: Results.....	98
6.1 Descriptive Results.....	98
6.2 Multivariate Results	106

6.2.1 Model 1	106
6.2.2 Gender Differences	113
6.3 Sensitivity Analysis Results	115
6.3.1 Missing Data	115
6.3.2 Measurement of Risk Tolerance	119
6.3.3 Timing of Variable Measurement	121
Chapter 7: Discussion	123
7.1 Interpretation and Evaluation of Results	123
7.2 Limitations	126
7.3 Implications and Future Research	127
7.4 Conclusion	130
References	132
Appendix A: Age-Adjusted Risk Tolerance Computation	147
Appendix B: Proportion and Means Tests in Table 6.1	152
Appendix C: Average Marginal Effects	155

List of Tables

Table 5.1. Variable Names and Coding.....	73
Table 5.2. Variable Names and Coding for Model 1.IMP.....	94
Table 5.3. Variable Names and Coding for Models 2 – 5	96
Table 6.1. Descriptive Statistics of Complete Cases	103
Table 6.2. Relative Risk Tolerance at Interview Date	105
Table 6.3. Relative Risk Tolerance at Interview Date from Light and Ahn (2010)	105
Table 6.4. Model 1: Logistic Regression of Higher Education Enrollment by Age 20..	110
Table 6.5. Model 1: Average Marginal Effects	112
Table 6.6. Mean Predicted Probability of Enrollment by Quartile of RRT	113
Table 6.7. Mean Predicted Probability of Enrollment by RRT Category.....	113
Table 6.8. Likelihood Ratio Test: Pooled versus Separate Logits for Men and Women	114
Table 6.9. Gender Differences and Key Parameters.....	115
Table 6.10. Missing Data Analysis: Descriptive Statistics By Selected.....	117
Table 6.11. Sensitivity to Missing Data Using Imputation.....	119
Table 6.12. Sensitivity to Measurement of Risk Tolerance.....	120
Table 6.13. Likelihood Ratio Test: Pooled versus Separate Logits for Respondents Younger than Age 16 and Respondents Age 16 or Older.....	121
Table 6.14. Sensitivity to Timing of Variable Measurement.....	122
Table A.1. Upper and Lower Bounds for Risk Tolerance	148

Table A.2. Light and Ahn's (2010) Estimates of Risk Preference Parameters.....	149
--	-----

List of Figures

Figure 3.1 (Sociological) Relative Risk Aversion Theory.....	48
Figure 3.2. Student College Choice Process	52
Figure 3.3. Conceptual Model of College Choice	55
Figure 3.4. Financial Nexus Model.....	58
Figure 5.1. Map of U.S. Census Regions.....	80
Figure 5.2. Income Gamble Question Branching and Risk Categories	83

Chapter 1: Introduction

1.1 Background

Concern over the growing amount of student loan debt has dominated popular media outlets in recent years. Media sources (Harris, 2011) and researchers (Harrast, 2004; Kamenetz, 2006) have argued that education is increasingly becoming a bad investment due to the incredible growth in tuition rates and subsequent debt levels at graduation. University administrators are concerned that student wellness is negatively being affected by personal finance issues such as growing debt burdens (e.g., The U.T. System Student Debt Reduction Task Force, 2012). In recent surveys, many students have reported feeling financially stressed or that financial stressors are the most important source of stress (Heckman, Lim, & Montalto, 2013; Trombitas, 2012). Some economists and policymakers are simply concerned about the size of the student loan burden in the United States and students' ability to repay (Harrast, 2004).

In 2012, outstanding student loan debt surpassed credit card debt as the largest source of non-mortgage household debt (Federal Reserve Bank of New York, 2012a). Student loan delinquencies recently grabbed headlines due to the increasing proportion of students falling behind (Hamilton, 2013) and by surpassing the number of credit card delinquencies (Federal Reserve Bank of New York, 2012b). Nearly half of all outstanding student loan debt is in deferred status (Hardekopf, 2013).

Furthermore, many more households are now dealing with concerns about student loans. Fry (2012) reported that among households headed by someone under the age of 35, 40 percent have student loans. Headlines referring to the “student loan bubble,” “student loan debt crisis,” and “student loan horror stories” encourage consumers and policymakers to jump to conclusions regarding the severity of the situation. Some researchers suggest that student debt to income ratios are too high (Baum & O'Malley, 2003; Hira, Anderson, & Petersen, 2000), but others point out that the average student debt to income ratio has remained stable in recent years (Avery & Turner, 2012). What remains unclear is the extent to which students are making rational decisions or errors in their decision to pursue a college education.

While traditional economic theory assumes that individuals are perfectly rational and make optimal decisions based on perfect information, a large literature of positive research in psychology, economics, and finance suggests that this is not the case. Daniel Kahneman and Amos Tversky (1979) were among some of the first researchers to begin to incorporate cognitive psychology into the study of economic behavior. They proposed prospect theory as a way to explain the empirical evidence that people often behaved in ways that departed from traditional economic theory. This has led to the development of several streams of research under the names of behavioral economics, behavioral finance, and economic psychology. The general conclusion from this vast area of research is that individuals are subject to a number of cognitive limitations and biases that influence decision making (see Hastie & Dawes, 2010; Koehler & Harvey, 2008). One of the largest concerns about the student loan problem is that students are not making good

borrowing decisions, which leads to defaults on loan payments and other financial problems upon graduation.

Since the field of behavioral economics suggests that, in general, consumers are prone to make mistakes in risky decisions (e.g., see Ariely, 2008; Kahneman, 2011; Thaler & Sunstein, 2008), the fact that higher education choices have not been studied in great detail as a risky decision is somewhat surprising. This is even more surprising considering that there is empirical evidence that may suggest that there are problems in the decision making process among students. For example, many students are unable to estimate what they owe in student loans (Andruska, Hogarth, Fletcher, Forbes, & Wohlgemuth, 2012; Hira et al., 2000). This raises concerns that students do not properly understand the costs involved in pursuing higher education. Other research has shown that students overestimate the costs of attending higher education (Avery & Kane, 2004; Oreopoulos & Petronijevic, 2013), which would negatively impact the evaluation of a college degree in a traditional human capital investment model.

Furthermore, researchers have found that students are not matching their prospective careers with educational requirements (Goyette, 2008; Schneider & Stevenson, 2000). Given the substantial variation in earnings by field of study (Carnevale, Cheah, & Strohl, 2013), we would expect to see students making poor investment decisions without this important piece of information. Lastly, the prevalence of stress from personal finances (Heckman et al., 2013) also suggests a need to improve the decision making process so students are more comfortable with their education and financing decisions.

From the consumer's point of view, the question of whether or not to attend higher education is a complex decision. While many resources are available to help educate consumers about various forms of financial aid, including loans, grants, scholarships, and work-study, there are comparatively few resources available to help students determine how much to invest in higher education. Of the few resources available, they primarily focus on rules of thumb (Clark, 2009) or simple calculations based on financial need and expected family contribution (Kantrowitz, 2013). Although there is a growing trend towards promoting informed consumer decision making, such as the White House's College Scorecard initiative (College Affordability and Transparency Center, 2014), there remains a lack of understanding regarding how students perceive and handle the risk involved in higher education choices.

Consumers must make higher education choices based on imperfect information and uncertain outcomes regarding college completion and labor market outcomes post-graduation. Clearly, there is a need for consumer guidance in this area. As suggested by Avery and Turner (2012), specific, individualized assistance is likely needed for consumers. However, the current state of research leaves many questions regarding the ways in which consumers make decisions about higher education investments.

1.2 Purpose

Since prior research has shown the positive effects of decision aids on college choice satisfaction (Kmett, Arkes, & Jones, 1999), helping consumers to follow a rational decision making process should lead to good investment decisions and a number of possible positive benefits (e.g., increased satisfaction, decreased stress, etc.). The

motivating interest of the current study is to investigate whether consumers are making rational investment choices in higher education given the risk involved in the decision. Since risk plays a central role in the higher education decision making process, one way to define rationality, at a very fundamental level, is to posit that consumer risk perceptions and risk preferences should affect consumer educational choices. While this analysis will not provide the definitive answer on whether or not consumers are making rational choices in human capital investments, this analysis will provide a baseline measure of rational thinking by determining whether risk preferences influence the enrollment decision. Subsequently, the guiding research question is the following:

RQ: *Ceteris paribus*, what effect do consumer risk preferences have on higher education enrollment decisions?

Several assumptions are implicit in phrasing the research question in this way. First, the primary interest is in the decision making process as opposed to outcomes after college (e.g., wage rates, debt ratios, etc.). Second, higher education is assumed to be an investment in human capital and, as such, the motivating reason for pursuing higher education is the prospect of increased wages. Considering the college experience as a consumption good is certainly valid and consumers are probably willing to pay substantial sums of money to live the college experience. That is, there may be a substantial amount of direct utility to be gained from the college experience. Furthermore, there are a number of non-pecuniary benefits of attending higher education (e.g., networking opportunities, social status attainment, etc.). Although pursuing higher education does have consumption value and non-pecuniary benefits, the focus of this research is the monetary benefits of attending higher education. There are two primary

reasons for this focus: (1) the small literature regarding the consumption value of higher education does not provide guidance in terms of the importance of the consumption value of education relative to the pecuniary returns, and (2) understanding how consumers evaluate the monetary returns to higher education will provide a starting point for investigating the consumption value and non-pecuniary benefits.

1.3 Significance

While traditional normative economic approaches to human capital investment have provided useful tools to analyze the college financing decision, these models do not accurately portray the uncertainty involved (Dickson & Harmon, 2011). The traditional approaches have generally assumed perfect certainty, which is far from the reality that consumers face today. Other researchers seem to agree with this assessment. For example, Avery and Turner (2012) suggest that high school graduates need more individualized help, based on observable characteristics, in determining the amount they should borrow. Specifically, researchers suggest that high school graduates could consider variance in earnings between fields of study (Avery & Turner, 2012) and estimate the probability of degree completion (Avery & Turner, 2012; Oreopoulos & Petronijevic, 2013) when making the decision to invest in higher education.

The current research is unique in that higher education choices, rather than the outcomes of attending higher education, are emphasized as decision making under uncertainty. This approach evaluates *how* a decision should be made as opposed to *which* decision to make (Hastie & Dawes, 2010). Furthermore, to the author's knowledge, this is the first study to use measures of both subjective probabilities and time preferences in

conjunction with individual-specific risk preferences in estimating the likelihood of enrolling in higher education. Emphasizing the decision making process is a reasonable approach due to the risky nature of higher education decisions.

Given the uncertainty and risk involved in the decision to invest in higher education, it may be more productive to study the decision making process to determine how best to help consumers as opposed to focusing primarily on outcomes, such as post-graduation earnings or debt levels. For example, students cannot be blamed for taking loans and investing in higher education if unemployment soars as they graduate or if technological advances make their training obsolete. Circumstances such as these, in addition to a plethora of others, highlight the risky nature of such decisions.

In addition to the unique empirical model, this study makes a contribution to the field by providing an interdisciplinary review of research regarding higher education decision making. Since scholars come from a variety of backgrounds, including consumer sciences, economics, education, finance, and sociology, the literature tends to be somewhat segregated and challenging to synthesize. Therefore, the literature review and the review of the theory is a useful and original synthesis of research regarding consumer higher education choices.

The findings of this study will be useful to consumers, researchers, administrators, and policymakers who seek to understand and/or improve consumers' higher education enrollment decisions. The role that risk perceptions and preferences play in higher education choices are not well understood but these constructs have the potential to have a significant – and possibly negative – impact on consumer behavior. This research is a starting point for investigating these issues.

Chapter 2: Literature Review

In order to grasp our current knowledge about consumer decision making regarding higher education, several areas of literature must be examined. The relevant literature was identified by considering the various decision points and outcomes for a student who considers pursuing a college degree. Subsequently, this literature review is organized topically by the various decision points and considerations.

As a beginning point, we must recognize the fact that the outcomes to higher education are inherently unknown; that is, college education is a risky decision. Therefore, Section 2.1 provides a general overview of the literature regarding consumer risk preferences, and Section 2.2 reviews the literature specific to risk and higher education. Since higher education is both expensive and voluntary, Section 2.3 reviews the literature pertaining to higher education enrollment. Furthermore, since many students begin college but are unable to persist to graduation, Section 2.4 reviews the persistence literature. Lastly, there is an immense literature regarding the returns to education. While the current study will not directly contribute to this literature, Section 2.5 reviews this research primarily as an empirical consideration due to the important effect that the returns to education are believed to have on the decision to pursue higher education. Section 2.6 concludes by summarizing the relevant gaps in the literature and by specifying how this study will advance our understanding of higher education decisions.

2.1 Risk and Consumer Risk Preferences

As a necessary foundation for decision making research, the concepts of risk and consumer risk preferences are reviewed in the following discussion. Although the terms are popularly used as synonyms for one another, uncertainty and risk should be carefully defined. While both uncertainty and risk involve the occurrence of future events, the distinction lies in whether the probability distribution of the event is known or unknown (Knight, 1921). Risk involves future events in which the probability distribution of occurrence is known while uncertainty involves future events in which the probability distribution is unknown (Knight, 1921). In consideration of the decision to invest in higher education, the return on the investment (i.e., was it worth it?) is dependent upon several future events: graduation, labor market outcomes, affordability of debt payments, future macroeconomic conditions, and a number of other unknowns (Mazza, 2011). Since the primary focus of the current study is consumer decision making, the distinction between risk and uncertainty is vitally important. Decision makers can do little about true uncertainty, but managing risk is possible through decision making strategies, such as expected utility analysis. Risk, therefore, is the relevant consideration for the current study.

Consumer attitudes towards and willingness to accept risk has been the subject of a large literature in a variety of contexts. Two important concepts in this literature are the risk premium and risk aversion. The risk premium (R) is the amount of money that a consumer requires in order to be indifferent ($\sim^*$) between the payoff of a risky choice (a) and the same payoff of a certain choice ($E(a)$) (Chavas, 2004). Therefore, a positive risk

premium indicates that the consumer is risk averse while a negative risk premium indicates that the consumer is risk-loving. A risk premium equal to zero implies that the consumer is indifferent to risk (i.e., risk neutral). The majority of risky decision making research reviewed has confirmed risk averse preferences (e.g., Hanna, Gutter, & Fan, 2001), and the existence of insurance markets implies that consumers are risk averse (Chavas, 2004). Furthermore, research has consistently shown that risk preferences differ systematically between males and females (e.g., Grable, 2000; Grable & Lytton, 1998; Neelakantan, 2010; Sung & Hanna, 1996); therefore, the standard practice in applied research is to separate men and women for empirical analysis (e.g., Light & Ahn, 2010).

The stability of risk preferences, both over time and between contexts, is an important consideration when analyzing risky decisions. Recent research suggests that, contrary to previous assumptions, risk preferences are time-variant (Ahn, 2010; Light & Ahn, 2010; Sahm, 2012). Specifically, consumers generally become less willing to take risk (i.e., more risk averse) as they age (Ahn, 2010; Light & Ahn, 2010) and risk preferences are influenced by macroeconomic conditions (Sahm, 2012). However, Sahm (2012) estimates that 70 percent of the systematic variation in risk preferences is a result of individual-specific time-invariant heterogeneity.

Scholars have also considered whether or not the specific context of the risky choice influences consumers' willingness to take risk. For example, does a consumer's risk preference regarding financial investment decisions differ from his or her willingness to take risks in health behaviors, such as smoking? Weber, Blais, and Betz (2002) found that willingness to take risk is highly dependent upon the decision context; however, they concluded that the differences in willingness to take risk were mostly a result of

differences between the expected benefits and perceived risk – not the underlying risk preference.

In a recent study, Dohmen et al. (2011) used an observational study and an experimental study to analyze consumer risk preferences. They found that risk attitudes are strongly, but not perfectly, correlated across decisions contexts. The Dohmen et al. (2011) study also provides validation for survey elicitation of consumer risk preferences. Specifically, they used an experimental design to validate survey responses to questions about their willingness to take risks. They found that a general measure on a 10-point Likert scale asking “How willing are you to take risks in general?” is the best predictor of risk preferences across contexts. However, they suggested that a risk question that incorporates the decision context into the question is the best predictor of behavior in any given context (Dohmen et al., 2011).

This section has reviewed important concepts that will be key considerations in the current analysis of risky human capital decisions. For the remainder of this study, risk will refer to future events with known probabilities, as consumers can be expected to make use of this information when making decisions. Furthermore, key concepts such as risk aversion and the findings regarding the stability of risk preferences will be critical considerations in this study.

2.2 Risk and Education Decisions

A common criticism of traditional economic theory is that standard models assume perfect information and certainty. While there are large literatures in economics and psychology concerning a variety of risky decisions, the literature has only recently returned to the issue of risk in college-going decisions and the returns to education. Aside from a handful of early studies of risky human capital (e.g., Eaton & Rosen, 1980; Kodde, 1986; Levhari & Weiss, 1974; Williams, 1979), it has only been in the last five or ten years that interest in this area has reemerged. There remains, therefore, a lack of empirical work in this area (Dickson & Harmon, 2011).

Risk arises in human capital decisions from the inability to know the outcomes of several future events: degree completion (Hartog & Diaz-Serrano, 2007; Mazza, 2011), future labor market supply and demand (Kodde, 1986), individual longevity which affects the duration of earnings realization (Kodde, 1986), location in earnings distribution (Hartog & Diaz-Serrano, 2007; Kodde, 1986; Levhari & Weiss, 1974; Mazza, 2011), and unemployment rates (Carnevale et al., 2013; Fossen & Glocker, 2011). Research has pointed out the need to distinguish future risk from private information (Mazza, 2011). That is, it is possible that the potential student acts upon information regarding productivity that is not available to the researcher. Several researchers (Hartog & Diaz-Serrano, 2007; Kodde, 1986) have suggested that the individual's own knowledge of this is imperfect at best, likely contributing to the uncertainty involved in the investment decisions.

To date, the majority of the research regarding risk in education decisions has focused on labor market outcomes; specifically, the uncertainty of future income as a

result of increased human capital. Furthermore, the expected relationship between risk and human capital accumulation is often theoretical in nature with limited empirical support. Levhari and Weiss (1974) were the first to examine the role of uncertainty in human capital decisions. They developed a simple two-period theoretical model and concluded that the effect of risk on human capital investment is generally ambiguous (Levhari & Weiss, 1974). Kodde (1986) presented a similar model to Levhari and Weiss (1974) but added an empirical investigation using a sample of high school students. He concluded that the effect of increasing risk cannot be signed unambiguously and that observed behavior seems to deviate from that which is expected (Kodde, 1986). More recent research found that attitudes towards risk were important predictors of the decision to attend higher education. Specifically, the higher the risk involved, the lower the investment in higher education; however decreasing risk aversion reduces the impact of risk on investment (Hartog & Diaz-Serrano, 2007).

Mazza (2011) found that uncertainty plays an important role in education choices, especially choice of major. He found that individuals pursue higher education in an attempt to minimize uncertainty in addition to maximizing returns. Interestingly, Mazza (2011) also found that the selection of humanities majors was not influenced by risk or expected returns. Several researchers have been interested in the effect of taxes on human capital risk and decision making. Eaton and Rosen (1980) found that taxes affect the riskiness of human capital and the willingness of consumers to assume the risk. Fossen and Glocker (2011) estimated a microeconometric model of the decision to enroll in higher education while accounting for three types of risk (dropout, unemployment, and earnings variation) and simulated tax changes on the likelihood of enrollment in

Germany. They found that higher risk-adjusted wages increase the likelihood of enrollment while greater variance of net wages for college graduates decrease the likelihood of enrollment (Fossen & Glocker, 2011). Other researchers have argued that the risk of not completing college is important enough that student loans should be insured (i.e., insurance products that pay student loans back in case of not completing a degree) to encourage enrollment (Chatterjee & Ionescu, 2012).

Hartog and Diaz-Serrano (2007) showed that increased earnings risk causes a decrease in post-secondary attendance. In their empirical model, they control for heterogeneous risk attitudes by using expenditures on lottery tickets as proxy for risk attitudes. Fossen and Glocker (2011) found that higher expected returns to education increase the probability of enrollment while greater uncertainty decreases enrollment probability. However, they used a structural econometric model to estimate a constant relative risk aversion parameter – i.e., risk preferences were estimated and not based on consumer responses (Fossen & Glocker, 2011, p. 695).

There are several studies that use consumer risk attitudes as predictors of higher education attainment. Belzil and Leonardi (2007) analyzed an Italian dataset to determine the effect of risk preferences on higher education attainment. They found only a modest effect of risk preferences, measured by a hypothetical lottery question, on schooling decisions. In a working paper, Brown, Ortiz, and Taylor (2006) analyzed the 1996 Panel Study of Income Dynamics to determine the effect of consumer risk preferences on human capital accumulation and the effect of parental risk tolerance on consumer human capital accumulation. They found that individual levels and parental levels of risk aversion were inversely related to education attainment; that is, increased levels of risk

aversion lead to lower education attainment (Brown et al., 2006). Their results are dependent upon time-invariant risk attitudes since the attitude towards risk was measured after the human capital decisions were made.

In another working paper, Belzil and Leonardi (2009) analyzed an Italian dataset by using a dynamic, discrete choice model to investigate the relationship between individual-specific time-invariant risk aversion and the likelihood of enrolling in higher education. They found that risk aversion, measured by a hypothetical lottery question, acts as a deterrent to higher education attainment (i.e., a negative relationship) and suggest the need to disentangle subjective probabilities from risk attitudes (Belzil & Leonardi, 2009).

Recently, another strand of literature has begun to investigate the relationship between risk and return in human capital investments (Andini, 2009; Athanassiadis, 2011; Cheng, 2013; Christiansen, Joensen, & Nielsen, 2007; Koerselman & Uusitalo, 2013; Mazza, van Ophem, & Hartog, 2013; Palacios-Huerta, 2003). As of yet, there is not a general consensus regarding this relationship. While this literature has implications in a number of areas, for the purposes of this study, it is sufficient to assume that consumers expect the relationship between risk and return in human capital investments to be positive.

This section has highlighted the fact that investment in human capital is indeed a risky decision and that the collective understanding of the influence of risk and risk preferences on educational decisions is not yet complete. A small number of studies have empirically investigated the effect of risk preferences on education decisions. However, there are only a few studies that have used direct measures of consumer risk preferences

when analyzing education decisions. To the author's knowledge, no human capital studies have examined risk preferences while controlling for time preferences and subjective probabilities.

2.3 College-Going Behavior

Prior economic research has analyzed the decision to enroll in higher education based on human capital theory and, specifically, the expected net present value of the investment (Catsiapis, 1987). Human capital theory, discussed in detail in Chapter 3, suggests that, in an effort to maximize satisfaction, individuals invest in human capital (i.e. enroll and persist in higher education) until the marginal benefits equal the marginal costs (Becker, 1962, 1964). Therefore, when regarded as an investment opportunity, estimating the returns to education has important economic implications for individuals and policymakers, and the return to a college degree should be a primary incentive to pursue higher education. This has resulted in a vast economic literature focused on estimating the returns to education, discussed in Section 2.5. For now, we will consider broad factors that have been associated with enrollment decisions.

Consistent with human capital expectations, Catsiapis (1987) found that greater financial incentives lead to greater likelihood of enrollment and that grants and scholarships had the largest positive effect on enrollment. While higher expected wages may encourage investment in higher education, it is also true that reduced (increased) costs should increase (decrease) investments. Research has confirmed the importance of financial aid in encouraging student enrollment in higher education. Specifically, Dynarski (2003) found that grant aid increases the likelihood of college attendance by 3.6

percentage points for each \$1,000 of grant aid. Research has also shown that higher tuition prices decrease enrollment rates (Card & Lemieux, 2001; Hemelt & Marcotte, 2011). State-based merit scholarships, such as the Georgia HOPE scholarship, have also been associated with increased likelihood of attendance (Dynarski, 2004). Since foregone earnings are a standard feature of the costs of higher education in human capital models, it is worth noting that research has found that higher regional unemployment (i.e., lower costs) encourages college-going behavior (Hartog & Diaz-Serrano, 2007).

Higher education researchers have suggested that the interaction between expectations regarding the costs and benefits of education play an important role in the decision to enroll in higher education and subsequent persistence decisions (Paulsen & St. John, 1997; St. John, Paulsen, & Starkey, 1996). Although this has generated a large amount of research into the financial considerations involved in attending and persisting in higher education (e.g., Paulsen & St. John, 2002; St. John, Paulsen, & Carter, 2005), researchers have found evidence that students typically overestimate both the costs of higher education and the expected benefits (i.e., wage premiums) of attaining a college degree (Avery & Kane, 2004).

Dating back to the 1970's, the literature has shown that a student's ability and family background or socioeconomic status (SES) have been very important predictors of the decision to enroll in higher education (Christensen, Melder, & Weisbrod, 1975). Perna and Titus (2004) found that higher education enrollment decisions differ by levels of socio-economic status (SES). Belzil and Hansen (2003) estimated that family background, which includes parental education, income, and family structure, accounts

for nearly 70 percent of the variation in education attainment. There are a variety of explanations for this finding, several of which are reviewed here.

One explanation is that budget constraints and willingness to borrow differ between SES levels. Given the current financial aid landscape in the United States (i.e., heavy reliance on student loans), researchers are worried that being from a low SES background negatively affects students' willingness to borrow. Findings from the U.K. indicate that the effect of attitudes towards debt on enrollment varies by social class – students from lower SES families tend to be more debt averse than similar students from higher SES families (Callender & Jackson, 2005). Students from low SES families also tend to view higher education costs as debt rather than an investment opportunity (Callender & Jackson, 2008). For low-income students, tuition has been negatively correlated with enrollment (Lassila, 2011; Savoca, 1990) while total grants and institutional revenues were positively associated with enrollment (Lassila, 2011).

SES may also be important in shaping education aspirations. Research has shown that students who live with both biological parents (i.e., an intact family) are significantly more likely to apply, to be admitted, and to attend higher education than students who do not live with both biological parents (i.e., a disrupted family) (Lillard & Gerner, 1999). Additionally, students from intact families typically attend more selective institutions than students from disrupted families (Lillard & Gerner, 1999).

Parents who are college graduates also borrow greater amounts to pay for their child's college education compared to otherwise similar households (Cha, Weagley, & Reynolds, 2005). Furthermore, Avery and Hoxby (2004) found that among high-aptitude students, those whose parents attended more selective universities were less responsive to

cost variables in the choice of which university to attend. These findings may indicate that parents with college degrees (possibly indicating higher SES) tend to place more value and importance on attaining higher education. The importance of creating educational aspirations should not be overlooked as research has shown that exposure to educated professionals increases educational aspirations among poor children and that parental aspirations are positively correlated with children's educational attainment (Chiapa, Garrido, & Prina, 2012).

Another possible explanation for the effect of family background may relate to differences in earnings expectations for low-SES individuals. Prior research has shown that low-SES students have lower expectations of future wages (Delaney, Harmon, & Redmond, 2011), which, according to human capital theory, would lead to lower investments in education. Since sociological research emphasizes the importance of social and cultural capital (for a review, see Perna, 2006), these expectations may be based on perceptions of lower amounts of social and cultural capital.

Black and Hispanic students tend to enroll in higher education at much lower rates than White students (Perna, 2000). Blacks tend to be more sensitive to grants and tuition than Whites (St. John et al., 2005). Furthermore, due to income distribution differences between Blacks and Whites, St. John et al. (2005) suggest that loan-based financial aid policies favor White students. The differences in enrollment rates between races have also been explained by differences in capital available to students of different races (Perna, 2000). Perna (2000) suggests that lower rates of enrollment among Blacks and Hispanics are attributable to differences in the social and cultural capital available. The costs and benefits of higher education also differ between races; after controlling for

these differences, Blacks are more likely to enroll than Whites and Hispanics enroll at comparable rates as Whites (Perna, 2000). Perna and Titus (2004) also found that the state context, including state appropriations, tuition, financial aid to students, and primary and secondary education, was all important predictors of college enrollment.

Aside from systematic group differences and financial incentives, other factors have been put forth as having an important influence on the decision to pursue higher education. Some argue that the consumption value of the college experience is more important than the academic aspects for many students. Researchers have found that students are generally more willing to pay for non-academic aspects of college (Jacob, McCall, & Stange, 2013). Others have pointed out the improved marriage prospects as an important reason to attend college (see Oreopoulos & Salvanes, 2011). However, the research regarding non-pecuniary incentives is much more limited as it is difficult to measure the value of such benefits.

Aside from a few studies mentioned in Section 2.2, almost all studies of college enrollment have omitted the influence of risk and consumer risk preferences. While our knowledge regarding enrollment decisions has certainly improved over the last few decades, the literature has yet to fully explore the effect of risk on consumer higher education decisions.

2.4 Persistence

While college and university administrators have always had a financial stake in minimizing dropouts, the increased scrutiny from state and federal governments has heightened the interest in completion rates. Subsequently, research has analyzed

graduation rates in an effort to identify the factors that predict college graduation.

Although college student persistence has been operationalized in a variety of ways (see Letkiewicz et al., in press) and has examined a variety of schooling contexts, the focus here is on undergraduate degree completion. Student persistence (or stated negatively, dropout) has been the topic of much research in several fields of study, including higher education, economics, and sociology.

Pre-college academic performance (Titus, 2006a) and SES (Titus, 2006a, 2006b) have been found to have positive relationships with college completion. Since women tend to outperform men academically (DiPrete & Buchmann, 2013), gender also has an influence on student persistence. There is also strong evidence to suggest that cognitive ability is an important predictor of educational attainment (Belley & Lochner, 2007). As expected, research has shown that students who perform poorly (in terms of GPA) during the first year are significantly less likely to persist to graduation (Braunstein, McGrath, & Pescatrice, 2000). Torres, Gross, and Dadashova (2010) confirmed that GPA predicts persisting between terms and suggest that working during college may have a moderating effect on persistence through GPA. Furthermore, students who attend school on a full-time basis are significantly less likely to drop out compared to students who attend on a part-time basis (Robb, Moody, & Abdel-Ghany, 2011).

A number of financial factors have also been associated with student persistence. For example, Robb et al. (2011) found that students who had better financial behaviors were more likely to persist. Researchers have also been very interested in the effect of student financial aid on persistence rates, although the results are somewhat mixed. Braunstein et al. (2000) found that financial aid was not a significant predictor of

persistence. Similarly, Perna (1998) found a minimal direct effect of financial aid on completion rates; specifically, she found that loans were negatively associated with completion while work study was positively associated with completion. However, Ishitani and DesJardins (2002) and Dowd (2004) found that receiving financial aid was associated with lower dropout rates.

Recent findings have suggested that there is a non-linear relationship between loans and the likelihood of completing college (Dwyer, McCloud, & Hodson, 2012; Robb et al., 2011; Zhan, 2012). Specifically, smaller amounts of student loan debt lead to increased likelihoods of graduation while larger balances have the opposite effect, decreasing the likelihood of graduating (Dwyer et al., 2012). A number of studies have explored whether or not the type of financial aid received has an influence on persistence. St. John et al. (2005) found that the type of aid has disproportionate effects between racial groups. They suggested that loan-based aid tends to favor Whites while Blacks tend to be more sensitive to grants (St. John et al., 2005). Other research has shown that grants are much more effective in encouraging graduation rates than are student loans (Li, 2008).

Previous research has also suggested that the effects of debt vary by sex and economic background. Findings suggest that differences in labor market opportunities between men and women lead to men dropping out of college at lower debt levels than women (Dwyer, Hodson, & McCloud, 2013). The effect of debt also differs by the economic background of the student. Students from wealthy backgrounds tend to be less affected by student loan debt than students from poorer families (Dwyer et al., 2012). Family income has been shown to be an important determinant of degree attainment

(Dowd, 2004). In fact, recent research has shown that there has been “a dramatic increase in the effect of family income on college attendance rates” between the 1980s and early 2000s (Belley & Lochner, 2007, p. 2).

A number of scholars have pointed out the importance of the state context in determining completion rates. Given the highly diverse governance structure of higher education between states, there is substantial variation regarding incentives to attend college. For example, Georgia offers a full scholarship to all resident students who have a GPA of 3.0 or higher through the HOPE scholarship program. Dynarski (2008) analyzed the influence of state-level merit aid on the likelihood of college completion and found that state-level programs increase the proportion of educated persons from 27 percent to 30 percent. Additionally, Roksa (2010) argued that the distribution of enrollments between community colleges and four-year schools makes a difference in completion rates. She found that states with greater community college attendance had greater baccalaureate completion rates. Roksa’s explanation for this finding is that community colleges serve as a sorting mechanism which leads to better completion rates. Titus (2006a) found that state appropriations to higher education were associated with increased completion rates after controlling for individual and institutional characteristics.

Furthermore, the institutional context of the college or university that the consumer chooses is also thought to have an important effect on degree completion. This proposition is traced back to student persistence theories (discussed in Chapter 3) which posit that the institution’s commitment to student success is very important to completion rates (e.g., Bean, 1983; Tinto, 1975). Researchers have confirmed the importance of the

institutional environment. For example, in a study of students with a low socioeconomic status, Titus (2006b) found that the financial context (e.g., spending per full-time equivalent student) of the institution was important in determining completion.

Institutional selectivity also has an important effect on academic outcomes as researchers have confirmed a strong peer effect on academic performance; specifically, strong (weak) students tend to increase (decrease) the academic performance of peers (Winston & Zimmerman, 2004). Furthermore, living on campus and being involved on campus have been associated with increased completion rates (Titus, 2006a) and shorter time-to-degree (Letkiewicz et al., in press).

This section has reviewed a number of factors that have been associated with college student persistence. Although the current study will not build upon the persistence literature, the key risk under investigation is the likelihood of persisting to degree completion. This literature is, therefore, an important empirical consideration to the current study. Since this study is focused on the enrollment decision, only factors that are known to or can be reasonably predicted by the student at the time of the enrollment decision entered into the empirical model.

2.5 Returns to Education

The pecuniary returns to education are the most often cited rationale for the necessity and value of a college degree. Additionally, the returns to education have been of great interest to economists since education has a large impact on labor markets and has widespread public policy implications. Saying that this has resulted in an immense literature regarding the estimation of the returns to education is an understatement – there

is a vast body of work examining this effect. While the current study will not contribute to the existing returns to education literature, a sampling of the existing work is reviewed here since it will be especially helpful in developing an empirical model of consumer behavior in education decisions. That is, the expected returns to education are expected to play an important role in consumer decision making.

The majority of the estimates of the returns to higher education is based on the work of Jacob Mincer (Mincer, 1958, 1974). As explained by Heckman, Lochner, and Todd (2008), the elegance and simplicity of the Mincer earnings equations have made this approach the method of choice for estimating education returns for most of the literature. Mincer's (1974) model uses ordinary least squares (OLS) to estimate the returns to education, ρ , by estimating log wages (w) as a function of years of schooling (s) and labor market experience (x), as shown in (2.1) (Heckman et al., 2008).

$$\ln w(s, x) = \alpha + \rho s + \beta_0 x + \beta_1 x^2 + \varepsilon \quad (2.1)$$

Since the aim of the current study does not directly involve building upon the Mincer equation literature, readers are directed to other reviews that provide surveys of the literature of Mincer empirical work (e.g. Card, 1999; Psacharopoulos, 1994; Psacharopoulos & Patrinos, 2004).

One of the central issues in the returns to education literature is the issue of endogeneity. Since higher ability individuals are likely to pursue education in the first place, estimates using OLS will be biased and inconsistent. The difficulty in empirical work is that ability is very difficult to observe since there are many factors that may be associated with ability that would be impossible to observe in typical datasets. In a simple econometric model, the Mincer earnings function in (2.1) becomes (2.2) with the

unobserved ability (a) lumped into the composite error term (v) with the standard error term (ε) (i.e., $v_i = a_i + \varepsilon_i$).

$$\log(w_i) = \beta_0 + \beta_1 M + \beta_2 S + v_i \quad (2.2)$$

If ability (a) is assumed to be correlated with education (S), (i.e., $cov(a_i, S_i) \neq 0$), then the error term is correlated with the regressor, S , which violates the standard OLS assumptions (Wooldridge, 2009).

Arguably the most convincing approach that has been used to address the endogeneity issue in estimating the returns to education analyzes the earnings of twins. Since the primary issue is controlling for innate ability, studies of identical twins have effectively been used to help establish the causal effect of education on earnings (Checchi, 2006). As explained by Checchi (2006), in addition to twin studies and natural experiments (such as effects of educational policy reform), there are several proposed methods for reducing or eliminating the endogeneity problem. Instrumental variable (IV) estimation has been used to reduce the bias, although there is currently a substantial amount of debate about the strength and validity of various instruments, such as father's education. Despite these concerns, the weight of the evidence from this literature seems to suggest confirmation of the causal effect (Checchi, 2006).

Another potential source of endogeneity bias results from model misspecification (Wooldridge, 2009). If the true earnings function includes other factors, the omission of these variables in the econometric model yields biased estimates. Subsequently, researchers have documented important factors that, when left unaddressed, can significantly bias the estimates of the returns to education. These factors include each of the following: gender (Dougherty, 2005; Reisel, 2013), racial/ethnic identity

(Arcidiacono, Bayer, & Hizmo, 2010; Reisel, 2013), taxes (Heckman et al., 2008), risk (Mazza, 2011), in-school work experience (Light, 2001) and socio-economic status (SES) (Delaney et al., 2011). Therefore, these factors should be considered in empirical models that need to control for the expected returns to education.

As summarized by Oreopoulos and Salvanes (2011), the current consensus from the returns to education economic literature is that the monetary returns to one year of high school or college range from 7 to 12 percent, on average. Researchers have consistently confirmed the existence of the so-called “sheepskin effects” which refers to exceptionally large returns to education in the year that the degree is awarded (e.g., Heckman et al., 2008; Jaeger & Page, 1996). The data also suggests that the wage premium has been growing (Avery & Turner, 2012; Bound, Lovenheim, & Turner, 2006; Oreopoulos & Petronijevic, 2013).

Oreopoulos and Salvanes (2011) outlined a host of other non-pecuniary private benefits, including improved work environments, sense of accomplishment, autonomy, job security, opportunity for social interaction, and prestige. They also point out that increased education can lead to better marriage opportunities, improved health choices, and that the education experience itself may provide utility, concluding that the combined effect of pecuniary and non-pecuniary returns can be quite large (Oreopoulos & Salvanes, 2011). Those with college degrees also face much lower unemployment rates, although there is substantial variation between majors (Carnevale et al., 2013).

Despite concerns about selection effects, the overwhelming evidence from the literature regarding the returns to education suggests that there is a causal, positive relationship between schooling and earnings, lending credibility to human capital theory

as a useful theoretical framework. Furthermore, this literature has identified important sources of bias that need to be considered in empirical models in order to account for expected returns to education. Lastly, although there are clearly non-pecuniary benefits to higher education, the importance consumers give to various benefits of attending higher education is not clear.

2.6 Summary and Gap

This chapter has reviewed several areas of an immense literature regarding higher education choices that are relevant to the current study. Given the uncertainty involved in human capital decisions, consumer risk preferences should have an influence on enrollment decisions if consumers are rational. Therefore, one way to improve our understanding of the decision making process is to determine the effect that consumer risk preferences have on enrollment decisions. The literature has identified a number of factors associated with college enrollment and dropout but has generally ignored the effect of risk and consumer risk preferences on these decisions. While the pecuniary benefits of higher education are well-documented, increased student loan borrowing and growing concerns about defaults raise questions as to whether consumers are making good investment choices.

This survey of the literature has also identified notable opportunities for future research. Very few studies of higher education decisions have considered the influence of risk and risk preferences. With the exception of a few studies (Belzil & Leonardi, 2007, 2009; Brown et al., 2006), empirical models of higher education decisions have typically relied upon controlling for individual characteristics or have assumed homogenous time-

invariant risk preferences. Furthermore, to my knowledge, existing studies have not used subjective probability estimates in empirical models of risky human capital decisions. This is an important consideration because prior research has shown the importance of subjective probabilities in consumer risk perceptions (e.g., Weber et al., 2002) and has suggested the need to disentangle subjective probabilities from risk preferences (Belzil & Leonardi, 2009). Therefore, this study sought to fill this gap in the current literature by determining the *ceteris paribus* influence of consumer risk preferences on the decision to enroll in higher education.

Chapter 3: Review of Theoretical Approaches

Given the variety of theoretical frameworks that have been applied to higher education choices, especially between disciplines, a brief review and evaluation of the different approaches was needed in order to understand how researchers have approached the higher education decision from a theoretical standpoint. Although there are a few existing summaries of theoretical approaches to higher education decisions (e.g., Hossler, Braxton, & Coopersmith, 1989; Paulsen, 1990; Perna, 2006), an update is needed to reflect recent trends toward treating the higher education decision as a risky decision. While not meant to be a comprehensive review of each approach, this chapter serves as an introduction to the various conceptual frameworks that have been used in studying higher education choices.

This chapter is organized as follows. Section 3.1 reviews economic approaches to education decisions, including human capital theory, signaling theory, expected utility theory, and revisits human capital models that incorporate expected utility theory into the framework to account for risk. Section 3.2 reviews sociological approaches to higher education decisions, including social and cultural capital and relative risk aversion theory. Section 3.3 reviews theoretical approaches used by higher education researchers. Each section includes a summary of the approaches within each discipline and the

chapter concludes with an overall summary of the theoretical approaches used in higher education decision making research.

3.1 Economic Approaches

This section reviews the two most common approaches in economics to studying higher education choices: human capital theory (3.1.1) and signaling theory (3.1.2). Since investing in higher education is a risky choice, expected utility theory (3.1.3) and risky human capital models (3.1.4) are also reviewed. This section concludes with a brief discussion of the economic approaches (3.1.5).

3.1.1 Human Capital Theory

Human capital theory suggests that consumers will choose to invest in their own productivity or human capital in an effort to maximize utility or satisfaction (Becker, 1962, 1964). This investment may take a variety of forms. For example, individuals can enroll in formal education to further their knowledge and improve their skill set or they might spend money on healthcare services (Becker, 1964). Human capital theory has generated numerous theoretical models, ranging from simple, one-period models to complex, dynamic programming models. For simplicity, a one-period human capital model is reviewed here.

A simple one period model, although stylized and not intended to capture all of the complexities of the decision to invest in higher education, can be very powerful in framing the decision to invest in education. Following Bryant and Zick's (2006)

explanation, suppose consumers intend to maximize utility (U) and utility is a function of market goods (X), leisure (L),

$$U = u(X, L) \quad (3.1)$$

In a one period model, consumers face a time constraint where total time (T) can be spent in the labor market (M), in leisure (L), or in school (S), so that

$$T = M + L + S. \quad (3.2)$$

Consumers also face a budget constraint where total income (Y) is determined by the market wage (w) and non-wage income (V).

$$Y = wM + V. \quad (3.3)$$

Lifetime expenditures consist of market good expenses, the product of the quantity of X and the price p , and the product of time spent in school and the direct costs of schooling units (C). By the satiation assumption (i.e., more is better) and ignoring bequests motives (since there is only one period), individuals are expected to consume all of their resources in their lifetime. Therefore, by equating lifetime expenditures to lifetime income, the budget constraint is shown in (3.4),

$$pX + CS = wM + V. \quad (3.4)$$

In order to account for the time constraint, (3.2) is substituted for M in (3.4),

$$pX + CS = w(T - L - S) + V. \quad (3.5)$$

By distributing the w and rearranging, the lifetime constraint is

$$pX + wL + (w + C)S = wT + V, \quad (3.6)$$

which simply says that expenses ($pX + wL + (w + C)S$) must equal income ($wT + V$).

A number of assumptions need to be introduced at this point. Wage is assumed to be a function of time spent in school ($w = w(S)$). The marginal return to education is positive ($\frac{\partial w}{\partial S} = w_s > 0$) and diminishing ($\frac{\partial w_s}{\partial S} < 0$). Since consumers are expected to maximize utility, the objective is to maximize (3.1) subject to (3.6), that is

$$\max U = u(X, L) \quad s. t. \quad pX + wL + (w + C)S = wT + V. \quad (3.7)$$

Equation (3.7) can be solved using standard constrained optimization techniques by forming the Lagrangian function,

$$\mathcal{L} = u(X, L) - \lambda[pX + wL + (w + C)S - wT - V], \quad (3.8)$$

and taking the partial derivative of (3.8) with respect to S , which yields the following first-order condition:

$$\frac{\partial \mathcal{L}}{\partial S} = -\lambda[w_s L + (w + C) + w_s S - w_s T] = 0. \quad (3.9)$$

By simplifying, the first-order condition becomes

$$\frac{\partial \mathcal{L}}{\partial S} = w_s M - (w + C) = 0. \quad (3.10)$$

This simple model can reveal important information regarding consumer higher education decisions. According to (3.10), consumers should continue to invest in education until the marginal benefits ($w_s M$) equal the costs ($w + C$). Generally speaking, factors that increase the benefits should increase the demand for higher education while factors that increase the costs should decrease the demand for higher education. For example, increasing marginal returns to education or time spent in the labor market (e.g., longer work life expectancy) would increase the optimal schooling level while tuition increases or stronger labor market conditions (i.e., increases in foregone earnings) would decrease the optimal schooling level. This simple, one period

model is commonly generalized to multi-period investment models (e.g., Heckman et al., 2008) in which utility is a function of lifetime consumption or income.

3.1.2 Signaling Theory

Spence (1973) offered an alternative to human capital theory in explaining the decision to pursue higher education. In focusing on labor market outcomes, Spence posited that firms face great uncertainty during the hiring process. Employers do not know with certainty which candidates will prove to be productive employees. Furthermore, this information may not be immediately obvious after hiring an individual since there is often a learning curve associated with any new job function. Time is needed for an employer to have an accurate understanding of the productivity of the new employee. Therefore, Spence concluded that employer hiring decisions are investment decisions involving uncertainty.

From the job candidate's point of view, each individual has observable characteristics that may provide information that helps to reduce the employer's uncertainty in hiring decisions. Some of these characteristics are fixed, such as gender, race: while others are alterable, such as education. Spence referred to the alterable characteristics as signals. He suggested that individuals seek higher education as a way to signal their productivity to firms. Therefore, it is expected that individuals would be willing to acquire signals as long as two conditions hold: (1) employers are believed to understand the signal as providing information about productivity and (2) acquiring the signal is cost effective (Bryant & Zick, 2006). Following Checchi's (2006, pp. 176-179) approach, it is useful to construct a simple signaling model in order to illustrate the

conditions under which workers may be motivated to acquire a signal (i.e., formal education).

In a stylized economy, suppose there are only two types of workers: productive and unproductive. Productivity is intended to capture natural ability or talent levels. Productive workers are endowed with productivity P_2 and unproductive workers are endowed with productivity P_1 ($P_2 > P_1$). Spence (1973) suggested that previous experiences of employers are used to inform predictions about worker type. Therefore, assume that employers are generally aware that a fraction of all workers in the economy, η , possess P_2 while the rest $(1 - \eta)$ possess P_1 . If no other characteristics are available upon which to make an informed decision, the employer's best course of action is to assume that a job candidate is drawn at random from the population and that productivity is expected to be

$$E(P) = \bar{P} = \eta P_2 + (1 - \eta) P_1 . \quad (3.11)$$

Therefore, supposing marginal revenue of talent to be ϕ , the employer should offer the same wage, $\bar{w}$, to all potential employees,

$$\bar{w} = \phi \bar{P} = \phi [\eta P_2 + (1 - \eta) P_1] . \quad (3.12)$$

In this economy, productive employees are underpaid since they should receive a high wage (w_2) given their productivity ($\bar{w} < w_2 = \phi P_2$). Unproductive employees, on the other hand, are overpaid since they should receive a low wage (w_1) given their productivity ($\bar{w} > w_1 = \phi P_1$). This mismatch occurs because of the lack of observable characteristics that would allow employers to distinguish worker type. Thus, if a signal could be acquired that would help alleviate the information asymmetry, productive

workers would be motivated to acquire the signal in order to receive the higher wage (w_2).

Education has been posited to be a signal that can reduce the information asymmetry. Assuming employers recognize education to be a valid signal, they will offer a wage rate corresponding to level of education completed, such that

$$w_i = \beta(S_i) * S_i, \beta' > 0, i = 1, 2. \quad (3.13)$$

β is the marginal return to education and is a function of education level (S) so that the marginal return for education level 2 (S_2) is greater than the marginal return for education level 1 (S_1). In this model, individual preferences, V_i , can be represented by

$$V_i = w_i(S_i) - \gamma(S_i, P_i), \gamma_s > 0, \gamma_p < 0, i = 1, 2. \quad (3.14)$$

where $\gamma(S_i, P_i)$ is the direct and indirect costs of education, which increases with level of education and decreases with productivity.

A necessary condition for a valid signal is that the costs of education are negatively correlated with the endowed productivity level (Checchi, 2006). If this were not the case, unproductive workers would simply imitate the behavior of productive workers and acquire the signal, thus rendering the signal useless. In order to motivate signal acquisition, the wage increase associated with the higher signal must exceed the costs of obtaining the signal. Therefore, unproductive workers do not benefit from acquiring the signal as long as $w_2 - \gamma(S_2, P_1) \leq w_1 - \gamma(S_1, P_1)$. Productive workers will benefit from acquiring the signal as long as $w_2 - \gamma(S_2, P_2) > w_1 - \gamma(S_1, P_2)$. By rewriting these inequalities in terms of $w_2 - w_1$ and combining, the following condition is derived:

$$\gamma(S_2, P_1) - \gamma(S_1, P_1) \geq w_2 - w_1 > \gamma(S_2, P_2) - \gamma(S_1, P_2). \quad (3.14)$$

This condition implies that the difference in expected wages is the key requirement for the consumer considering higher education. The expected wage increase has to be greater than the costs associated with acquiring the signal. Thus, signaling models imply that the higher expected wages are the reason for the acquisition of the signal and education does not necessarily improve productivity.

3.1.3 Expected Utility Theory

A number of frameworks have been proposed from scholars from economics and psychology to either prescribe or describe risky decision making. This section focuses on expected utility theory as the framework of choice for prescriptive (i.e., normative) risky decision making. Although alternate theories, such as prospect theory (Kahneman & Tversky, 1979), may describe actual consumer behavior more accurately, competing frameworks have yet to achieve the prescriptive status that the expected utility framework holds.

Without question, the most widely accepted theory of rational choice is John von Neumann and Oskar Morgenstern's (1944) expected utility theory (Chavas, 2004; Hastie & Dawes, 2010). This normative theory rigorously describes rational decision making through the use of an axiomatic method. Von Neumann and Morgenstern proved that if rational decision making follows certain properties (the axioms), it is possible to assign numerical values and probabilities to different outcomes (i.e., mathematical expectation). This allows an ordering of the preferred choices based on the expected utility of each outcome. While many of the standard textbooks in economics (e.g., Kreps, 1990; Mas-Colell, Whinston, & Green, 1995; Varian, 1992) and judgment and decision making

provide coverage of the expected utility framework (e.g., Hastie & Dawes, 2010; Koehler & Harvey, 2008) the explanation presented here follows the concise explanation presented in Chavas (2004).

The expected utility hypothesis states that consumers faced with a risky decision (a), and preferences represented by a utility function, $U(a)$, choose to maximize the expected utility $EU(a)$, where E is the expectation operator (Chavas, 2004). Among two alternatives (a_1 and a_2), the consumer will choose a_1 (i.e., a_1 is preferred ($\geq^*$) to a_2) if and only if the expected utility of a_1 is greater than the expected utility of a_2 ; that is,

$$a_1 \geq^* a_2 \text{ if and only if } EU(a_1) \geq EU(a_2). \quad (3.15)$$

This hypothesis holds under the following axioms of consumer preferences: (1) Ordering and transitivity, (2) independence, and (3) continuity.

Axiom 1: Consumer preferences are ordered and transitive.

Ordered preferences means that, for any random variables a_1 and a_2 , one of the following must hold: the consumer

- (1) prefers a_1 to a_2 ($a_1 >^* a_2$),
- (2) prefers a_2 to a_1 ($a_2 >^* a_1$), or
- (3) is indifferent between a_1 and a_2 ($a_1 \sim^* a_2$).

Consumer preferences must also be transitive; that is, considering three random variables ($a_i, i = 1, 2, 3$), if a_1 is preferred to a_2 and a_2 is preferred to a_3 , then a_1 is preferred to a_3 .

Axiom 2: Consumer preferences are independent.

Independent consumer preferences imply the following relationship. For any random variable, a_i ($i = 1, 2, 3$), and any probability θ ($0 < \theta < 1$), then $a_1 >^* a_2$ if

and only if $[\theta a_1 + (1 - \theta)a_3 >^* \theta a_2 + (1 - \theta)a_3]$. That is, the consumer's preference between a_1 and a_2 does not depend on a_3 .

Axiom 3: Consumer preferences are continuous.

Continuous preferences mean that for any random variable, a_i ($i = 1, 2, 3$) where $a_1 <^* a_3 <^* a_2$, there exist numbers β and γ (where β and γ are between zero and one), such that a sufficiently small change in probabilities will not affect a strict preference. That is, $a_3 <^* [\beta a_2 + (1 - \beta)a_3]$ and $a_3 >^* [\gamma a_2 + (1 - \gamma)a_3]$.

Although two additional technical assumptions are needed in order to guarantee that the expected utility function is measureable (see Chavas, 2004), the three axioms shown above have received the attention of researchers (e.g., Kahneman & Tversky, 1979). The reason for this is that if human behavior is inconsistent with that which is predicted by the expected utility theory, it must be because one of the axioms have been violated (Chavas, 2004).

Another useful feature of expected utility theory is that it has led to a large literature regarding the nature of risk preferences. Arrow (1971) and Pratt (1964) showed that, given a utility function for money ($U(x)$), the functions $r(x) = -u''(x)/u'(x)$ and $r^*(x) = xr(x)$ can be interpreted as measures of local risk aversion (Pratt, 1964). The first function, $r(x)$, is known as the Arrow-Pratt coefficient of absolute risk aversion and the second, $r^*(x)$, is the Arrow-Pratt coefficient of relative risk aversion. Both measures can have special properties in which consumers exhibit constant (increasing, or decreasing) absolute (relative) risk aversion (see Chavas, 2004). For the purposes of this discussion, it is sufficient to note that relative risk aversion measures are quite common

in empirical literature since they have the convenient feature of being independent of unit of measurement (e.g., dollars, euros, etc.) (Chavas, 2004, p. 46).

3.1.4 Risky Human Capital Models

This section revisits alternate models of human capital theory in which the expected utility framework is applied. Although risky human capital models have been around for several decades, few researchers have really explored the effect of risk on human capital decisions. There are a number of other approaches to risky human capital theory that have been employed in the literature. For example, researchers have employed dynamic programming techniques to model risky human capital decisions (Williams, 1979). Other researches have expanded human capital models to include real options (Bilkic, Gries, & Pilichowski, 2012; Hogan & Walker, 2007; Jacobs, 2007). The real option approach is quite interesting in that these models highlight further characteristics of the decision to invest in higher education. Jacobs (2007) points out that the decision to invest in risky human capital is irreversible (i.e., consumers cannot recover the time spent in education) and that the decision can be delayed. There is, therefore, an option value to waiting to invest in higher education which increases required returns (Jacobs, 2007). Currently, real option frameworks treat the decision maker as risk-neutral although there are methods for incorporating risk preferences into the consideration (e.g., Kulatilaka, 1995). For the purposes of this study, a basic, risky human capital model is sufficient to show that consumer risk preferences and subjective probabilities of completion should have an influence on the decision to enroll in higher education.

Most approaches to risky human capital decisions incorporate expected utility theory into traditional human capital frameworks, thus accounting for risk preferences. The simplest dynamic models of risky human capital acquisition involve two period models of consumption, schooling choices, and wage uncertainty (e.g., Kodde, 1986; Levhari & Weiss, 1974). Many researchers have also formulated multi-period dynamic models of lifetime income to model human capital acquisition (e.g., Hartog & Diaz-Serrano, 2007; Heckman et al., 2008). This section begins with a review of the first human capital model involving risk in which the expected utility framework was incorporated into traditional human capital models.

Levhari and Weiss (1974) were the first to articulate a basic human capital model including risk. Their two-period framework models utility as a function of present (c_0) and future consumption(c_1).

$$U = u(c_0, c_1) \quad (3.16)$$

Consumers may spend time either in the labor market or in human capital investment (i.e., formal schooling). Let λ represent the proportion of total time spent in formal schooling and $1 - \lambda$ represent the time spent in the labor market. Future income, y_1 , is a function of human capital investment (λ) and the future state of the world (μ), which is assumed to be unknown.

$$y_1 = f(\lambda, \mu) \quad (3.17)$$

The consumer's objective is to maximize expected value of the lifetime utility of consumption (V) by choosing the amount of consumption in Period 1 and investment in human capital.

$$\max_{c_0, \lambda} V = E[u(c_0, c_1)] \quad (3.18)$$

Consumption is constrained by total lifetime wealth and is represented by

$$c_1 = (A + (1 - \lambda)y_0 - c_0)(1 + r) + y_1 \quad (3.19)$$

where A represents initial wealth, and r is the market rate of interest.

Maximizing (3.18) subject to (3.19) yields the following first-order conditions:

$$V_{c_0} = E \left[\frac{\partial u}{\partial c_0}(c_0, c_1) - (1 + r) \frac{\partial u}{\partial c_1}(c_0, c_1) \right] = 0 \quad (3.20)$$

$$V_\lambda = E \left[\frac{\partial u}{\partial c_1}(c_0, c_1)(-y_0(1 + r)) + f_\lambda(\lambda, \mu) \right] = 0 \quad (3.21)$$

Assuming that r is known, (3.20) can be rewritten to show that the ratio between the expected present and future marginal utilities is equal to the discount factor, so shifting utility between periods is not productive. That is,

$$E \left[\frac{\frac{\partial u}{\partial c_0}(c_0, c_1)}{\frac{\partial u}{\partial c_1}(c_0, c_1)} \right] = (1 + r). \quad (3.22)$$

In their analysis, Levhari and Weiss concluded that the effect of risk is generally ambiguous, except in the case of where risk increases with schooling – in that case, individual investment in higher education is discouraged.

Further work by Kodde (1986) corrected one of Levhari and Weiss's conclusions (that the expected rate of return is equal to the interest rate) but reached the same conclusion as Levhari and Weiss in the general case – that the effect of risk on the demand for education cannot be signed unambiguously. However, Kodde shows that in the special case of multiplicative earnings specifications, increasing risk implies decreasing demand for education. Assuming decreasing absolute risk aversion and increasing risk, increases in initial wealth increase the demand for education.

Furthermore, recent work by Hartog and Diaz-Serrano (2007) (reviewed in Chapter 4)

has confirmed the importance of risk preferences in the decision to enroll in higher education.

3.1.5 Summary of Economic Approaches

While human capital theory and signaling theory have generated a lively dialogue amongst researchers, the two theories are quite similar in their predictions. Both ultimately simplify the decision to an investment decision in which the costs are weighed against the benefits. Prior research has confirmed that students generally think about the higher education decision in terms of the cost and benefits (Perna, 2008). The main distinction between human capital and signaling is the direction of causation (Bryant & Zick, 2006). That is, human capital theory assumes schooling makes workers more productive while signaling theory assumes that more productive workers seek out higher education. While there is strong evidence of an increased wage premium upon completing a degree as opposed to another year of schooling (i.e., sheepskin effects) (e.g., Heckman et al., 2008; Jaeger & Page, 1996), recent research attributes as much as two-thirds of the college wage premium to productivity enhancement (Fang, 2006), suggesting that human capital theory provides a useful framework for exploring rational education choices. This section also reviewed the importance of expected utility theory as a prescriptive theory of risky decision making and the ways in which expected utility theory has been incorporated into risky human capital models to analyze risky education choices.

3.2 Sociological Approaches

Researchers who use sociological approaches to study higher education decision making are often concerned with differential education attainment between classes and the cause of education and wage gaps. This section reviews social and cultural capital (3.2.1), two important theoretical constructs often used in sociological research. In addition to social and cultural capital frameworks, this section also reviews relative risk aversion theory (3.2.2) which has been used to examine class differences in educational attainment. This section concludes with a brief discussion of sociological approaches (3.2.3).

3.2.1 Social and Cultural Capital

In her review, Perna (2006) showed that sociological research has often utilized social and cultural capital frameworks to analyze higher education choices and that this approach developed out of traditional status attainment models in sociology. This section provides a brief review of social capital and cultural capital and concludes with implications for research regarding higher education decisions.

Coleman (1988) sought to bring together sociological approaches, which emphasize the important influence the social context has on individual behavior, with economic approaches that emphasize individual agents who act rationally to maximize utility. He posited that social capital is a particular kind of resource available to economic actors. While taking many different forms, all types of social capital have two elements in common; namely that “they all consist of some aspect of social structures, and they facilitate certain actions of actors – whether persons or corporate actors – within the

structure” (Coleman, 1988, pp.S98). Social capital, therefore, is manifested within relationships between different actors as opposed to existing within an individual actor or process.

Coleman was especially interested in the way that social capital contributes to the creation of human capital. In his seminal piece, he showed the importance of social capital in the educational outcomes of children. Specifically, parental presence and attention to children was an important predictor of high school dropout rates among students (Coleman, 1988). Social capital has also been explained in slightly different ways by other researchers (e.g., Bourdieu, 1986; Putnam, 1995, 2001), however, each approach emphasizes the fact that social capital relies on societal relationships and the ability of an individual to draw upon the collective resources available within a particular group or social structure.

Drawing upon Bourdieu (1986) and Bourdieu and Passeron (1977), Perna defined cultural capital as “the system of attributes, such as language skills, cultural knowledge, and mannerisms that is derived, in part, from one’s parents and that defines an individual’s class status” (Perna, 2006, p. 111). Bourdieu’s (1986) discussion of cultural capital emphasizes the acquisition of various forms of cultural capital and the relationship it has to economic capital. His explanation highlights the time investment by the family in a child that builds the stock of cultural capital and ultimately influences both the acquisition of human capital and the return of investments in human capital. Cultural capital, therefore, is heavily dependent upon an individual’s upbringing and the amount of investment made by caregivers.

Social capital and cultural capital tend to be operationalized as family background or socioeconomic status in sociological research. Having access to educated persons (i.e., social capital), whether those are parents or others in the community, is thought to be very important in influencing educational aspirations. Cultural capital is also needed in order for an individual to succeed in each level of education. Both social capital and cultural capital are thought to be influential in the type of return that an individual may realize from investments in higher education. That is, given two individuals with equivalent amounts of human capital, the returns may differ based on variations in the stock of social and cultural capital.

3.2.2 Relative Risk Aversion Theory

Relative risk aversion (RRA) theory, a sociological theory of educational choice developed by Goldthorpe (1996) and Breen and Goldthorpe (1997), has also been used in studying higher education choices. Goldthorpe (1996), reflecting on the poor explanatory power of Marxist and liberal theories of class, sought to develop a theory that would explain persistent class differentials in educational attainment. By applying a weak form of rational action theory, Goldthorpe (1996) posited that individuals seek to achieve their goals among a set of alternative actions. Breen and Goldthorpe (1997) expanded Goldthorpe's (1996) original work by mathematically formalizing the theory.

Breen and Goldthorpe suggested that students and parents consider three factors when determining educational choices: (1) direct costs and foregone earnings, (2) likelihood of success if education is undertaken, and (3) the utility of each educational choice. Breen and Goldthorpe posited that the third factor, the utility of educational

outcomes, can be operationalized as the access to a particular social class provided by a given level of education. In a simple decision model, assume that the choice is to stay in school or leave and enter the labor market directly. For those that stay in school, there is risk that they will fail in higher education, which could be interpreted as a dropout risk. The three outcomes (pass, fail, or leave higher education) are associated with differing probabilities of entering three different social classes: the service class (S^*), the working class (W^*), and the underclass (U^*). In Figure 3.1, a simple decision tree is used to illustrate the choices and associated outcomes and probabilities.

Breen and Goldthorpe proposed that class differences persist in educational attainment due to three mechanisms: relative risk aversion, primary effects, and resource differences. In their use of the term relative risk aversion, Breen and Goldthorpe meant that families try to avoid downward social movement for their children so they seek to maximize the probability of access to a social class at least as good as their current class. That is, relative to his or her current social class, individuals are averse to downward movement. Primary effects refer to systematic variation in average ability (a) by social class; that is, $a_{U^*} < a_{W^*} < a_{S^*}$. Therefore, primary effects suggest that educational attainment is expected to be different between social classes due to differences in ability to succeed at higher education and differences in the subjective likelihood of completion. Lastly, available resources (r) to family and students to offset costs of higher education are expected to vary by social class, specifically $r_{U^*} < r_{W^*} < r_{S^*}$ (Breen & Goldthorpe, 1997).

Figure 3.1 (Sociological) Relative Risk Aversion Theory
Source: Adapted from Breen and Goldthorpe (1997, p. 280).

Breen and Goldthorpe also assumed that there is a social consensus regarding beliefs about the parameters in Figure 3.1 which can be specified as the following conditions:

1. Staying in school and passing increase the chances of entering the service class than does staying in school and failing or leaving. That is, $\alpha > \beta_1$ and $\alpha > \gamma_1$.

2. Staying in school and failing increases the chances of entering the underclass than does leaving. That is, $\gamma_1 + \gamma_2 > \beta_1 + \beta_1$.
3. Leaving school increases the chances of entering the working class relative to the service class. That is, $\frac{\gamma_2}{\gamma_1} > 1$.
4. Staying in school and passing increases the chances of entry into the service class compared to the working class. That is, $\alpha > 0.5$.

Given these conditions, Breen and Goldthorpe derived several important conclusions. If continuing in education is costless and there are no differences between classes in the subjective probability parameters (the Greek letters in Figure 3.1), then individuals from middle-class backgrounds (the service class) will more strongly prefer to stay in school compared to those from lower classes (the working class). This conclusion reflects the relative risk aversion effect. Breen and Goldthorpe also showed that the subjective probability of passing (π) captures individual differences in ability. Since the primary effects suggest that ability differs systematically between classes, the average value of π will be lower among individuals from the working class compared to the service class. Lastly, assuming that staying in higher education is costly, Breen and Goldthorpe showed that the proportion of individuals who stay in school from the service class will be greater than the proportion of working class due to the greater average level of resources in the service class compared to the working class.

3.2.3 Summary of Sociological Approaches

The sociological approaches reviewed in this section emphasize the importance of the social context in which individuals make decisions. Behavioral differences between classes are typically the focus of sociological research. Social capital and cultural capital have often been utilized to explain persistent differences in educational attainment between social classes and could also explain why the literature has found systematic differences in the returns to education based on gender and racial/ethnic background. Relative risk aversion theory stresses the importance of the likelihood of entering different social classes based on an individual's decision to enroll in higher education and whether or not he or she successfully passes.

3.3 Higher Education Approaches

In addition to the economic and sociological approaches to human capital acquisition, higher education researchers have formulated models of the college choice process and dropout decisions. This section reviews the Student College Choice Process (3.3.1), the Integrated Model of Student College Choice (3.3.2), and the Financial Nexus Model (3.3.3). This section concludes with a brief summary and discussion of the higher education approaches reviewed (3.3.4).

3.3.1 Student College Choice Process

Hossler and Gallagher (1987) suggested that college decision making follows a three-stage process. In the first stage, known as the predisposition stage, students begin to formulate attitudes and aspirations regarding higher education. This stage begins very

early in life, possibly as early as preschool, and continues throughout formal education. As expected, research has confirmed that socioeconomic status, ability, and attitudes of family and peers are influential during this stage (Hossler & Gallagher, 1987).

In the second stage, known as the search stage, students begin to proactively determine their educational choice set and alternative options (e.g., entering the labor market). During this stage, interaction between students and institutions increase as students seek out information regarding specific institutions and institutions seek out students via recruiting methods. Students formulate a choice set during this stage, which is the group of institutions that the student plans to apply to. Hossler and Gallagher noted that the actual process in which students settled on a choice set is diverse and that high-ability students tend to conduct more sophisticated searches.

In the third stage, the choice stage, students evaluate the choice set and institutions weigh in by recruiting activities and admissions and the stage concludes with the student's decision of which university to attend. This stage is characterized by increased interaction between the student and institution as financial aid packages are communicated and the institutional characteristics are evaluated by the student.

Hossler and Gallagher's model is presented in Figure 3.2. In this conceptual model, students may decide to forgo Stages 2 and 3 if they decide in Stage 1 that college is not a good option. The alternative of attending higher education is labeled "other" which can be interpreted to mean that the student opts for some other option (e.g., enters the labor market directly). Despite a positive predisposition to higher education, the student may decide to abandon the college choice process during Stage 2 and pursue an alternative option.

Figure 3.2. Student College Choice Process

Source: Created from Figures 1-4 in Hossler and Gallagher's (1987, pp. 208-217) conceptual model. boxes signify decision outcomes. Underlined text represents stage outcomes. Dark arrows represent decision points.

3.3.2 Integrated Model of Student College Choice

Perna (2006) sought to utilize the respective strengths of economic approaches (i.e., human capital theory) and sociological approaches (i.e., social and cultural capital) by proposing an integrated model of student college choice. Her conceptual model emphasizes the multi-layered context in which rational consumers make higher education decisions and stresses the importance of the influence that this context has on typical human capital decision frameworks.

At the center of Perna's model is a traditional human capital model in which expected benefits and costs play a central role in the decision to attend higher education. The expected benefits and costs are partially influenced by the student's academic preparation and available resources to pay for the costs of higher education (e.g., family resources, financial aid, etc.). Perna incorporated insights gained from sociological approaches into her conceptual model by positing that the contextual setting for the traditional human capital approach is very influential in higher education decisions. That is, the perceptions regarding important factors in the human capital framework are influenced by four layers of context: (1) habitus, (2) school and community, (3) higher education context, and (4) social, economic, and policy context.

Habitus refers to "an individual's internalized system of thoughts, beliefs, and perceptions that are acquired from the immediate environment" (Perna, 2006, p. 113). Habitus, therefore, is thought to be extremely influential in shaping preferences, attitudes, and educational aspirations. In Perna's model, factors included in this context include demographic characteristics, cultural capital, and social capital. The school and community context refers to the types of resources available and supports (barriers)

within the school or community that promote (inhibit) college choice (Perna, 2006). This may include teachers, counselors, and middle-class peers who can help disseminate information regarding college admission processes (Perna, 2006).

The higher education context refers to the important role that colleges and universities have in shaping higher education choices. Perna (2006) suggests three ways in which colleges and universities help shape this decision. First, institutions are a source of information about higher education options and may actively or passively recruit students. Second, the specific characteristics of institutions and the way in which this matches student interests and needs are posited to influence the college choice. Third, institutions can control the number and type of students they choose to admit.

Lastly, the social and economic policy context includes social demographic changes, economic conditions, and public policies (Perna, 2006). For example, labor market changes or educational policy changes might influence the evaluation of the costs and benefits of attending higher education. A simplified version of Perna's framework is presented in Figure 3.3.

Figure 3.3. Conceptual Model of College Choice
Source: Adapted from Perna (2006, p. 117).

3.3.3 Financial Nexus Model

The financial nexus model (Paulsen & St. John, 1997; St. John et al., 1996) sought to bring two strands of literature and theory together: the college choice literature and the persistence literature. The college choice literature referred to by Paulsen and St. John included Hossler and Gallagher's three-stage model in addition to human capital and status attainment models from sociology. The two most common approaches to

persistence, the Student Integration Model and the Student Attrition Model, are briefly mentioned in this discussion to provide context to the development of the financial nexus model.

Cabrera and Castaneda (1992) argued that the Student Integration Model and the Student Attrition Model are complementary and that a more complete understanding of persistence is achieved when the main propositions of each theory are combined. Tinto's (1975) theory of dropout, referred to as the Student Integration Model of Persistence posited that university completion is a function of the student's and the institution's goal commitment and academic and social integration. The student's educational aspirations are largely influenced by background characteristics such as family background, individual characteristics, and prior schooling experiences. Bean's (Bean, 1982, 1983; Bean & Metzner, 1985) Student Attrition Model adopted an employee turnover model and applied it to student decisions to leave higher education. This model also drew upon the Theory of Planned Behavior (Ajzen & Fishbein, 1977; Ajzen, Fishbein, Heilbroner, & Thurow, 1980; Fishbein & Ajzen, 1975), emphasizing the intent to leave as a predictor of actual attrition behavior. Bean's model posited that satisfaction decreased the intent to leave which in turn reduced the likelihood of dropping out.

Paulsen and St. John argued that focusing on one choice (either enrollment or persistence) while ignoring the other artificially splits related decisions. They posited that students form an implicit contract with institutes of higher education based on pre-matriculation expected benefits and costs. If this comparison of expected benefits and costs are positive, then the student enrolls. Failure to persist to degree completion is believed to be attributable to the interaction, or nexus, between the expectations and the

actual benefits and costs. In particular, the financial considerations are critical to the decisions of whether to enroll and remain in school. The financial nexus model is shown in Figure 3.4.

Figure 3.4. Financial Nexus Model.

Source: Slightly modified version of Figure 4.1 in Paulsen and St. John (1997, p. 66).

3.3.4 Summary of Higher Education Approaches

The higher education approaches reviewed here tend to be interdisciplinary in their approach to higher education decisions; they tend to incorporate key constructs from both economic and sociological perspectives. Perna's (2006) model affirms traditional human capital models but also acknowledges the importance of an individual's perception of the costs and benefits, which is influenced by the contextual setting of the decision. Both social capital and cultural capital constructs appear in her model as part of the context setting. The financial nexus model incorporates both traditional human capital considerations (such as cost and benefits) but also acknowledges the important influence that an individual's family background can have on the decision to enroll in higher education.

3.4 Summary

Given this review of the theoretical approaches to education decisions, several points are worth emphasizing. By far, the majority of approaches to the decision to invest in higher education make use of an investment approach in which the costs are compared to the benefits. Costs may be direct (e.g., tuition) or indirect (e.g., foregone earnings, effort, etc.) while benefits may be pecuniary (e.g., labor market returns to education) or non-pecuniary (e.g., social status attainment).

Much of the current literature ignores the risk involved in the pursuit of higher education, despite the fact that many of the theoretical approaches outlined in this chapter imply that higher education choices are risky decisions. In human capital models, risk is primarily operationalized as labor market risk (unemployment, wage premium, income

distribution, etc.) since the primary benefit to investing in human capital is the increased wages associated with greater productivity. Signaling theory emphasizes the risk to employers who are in need of a sorting mechanism that can reduce the information asymmetry regarding prospective employee productivity. In RRA theory, risk is operationalized as dropout risk and the associated likelihoods of social class entry conditioned on educational enrollment and outcome. In the financial nexus model, risk is involved since the pre-matriculation beliefs about costs and benefits may be different than the post-matriculation realizations of costs and benefits (i.e., the future costs and benefits are not known with certainty).

Despite the fact that risky human capital models, signaling models, RRA theory, and the financial nexus model all imply that risk is involved in the decision to pursue higher education, little is known about the influence of risk preferences on the decision to enroll in higher education. Although Paulsen and St. John (1997) point out that enrollment and persistence decisions are clearly related, this research is concerned with understanding the decision making process of consumers leading up to enrollment and especially the influence of risk perceptions and risk preferences. As such, focusing primarily on the enrollment decision is an appropriate place to start since trying to capture both enrollment and persistence greatly complicates the decision process and timing of decisions.

Since the economic approach to risky decision making has a long history of rigorous modeling of the effect of consumer risk preferences on optimal risky decisions, the current research will make use of a recently developed risky human capital model. Nearly all of the approaches reviewed here involve a comparison of the costs and benefits

so economic theory is a useful place to begin to explore consumer risk preferences and the decision to enroll in higher education.

Chapter 4: Theory

In order to isolate the effect of risk and subjective probabilities for the current study, a risky human capital model was adopted to guide the exploration of consumer risk preferences and the decision to enroll in higher education. This chapter outlines the theoretical model (Section 4.1), the empirical model (Section 4.2), and then concludes with the research hypothesis (Section 4.3).

4.1 Theoretical Model

A risky human capital model was chosen as the most appropriate theoretical framework for the current study. Hartog and Diaz-Serrano (2007) developed a multi-period human capital model in which the returns to education are uncertain. Their model is described as follows (see Hartog & Diaz-Serrano, 2007, pp. 5-7).

To begin, assume that consumers believe that annual earnings at time t are a function of realized schooling, s , and a random shock θ . That is,

$$Y_{st} = \theta_{st} Y_s. \quad (4.1)$$

Furthermore, by assuming θ is a lifetime shock to Y (i.e., the consumer does not know the wage when deciding but realizes a single lifetime shock), $\theta_{st} = \theta_s$. Assume that consumers generally interpret income to correspond to consumption so that income

uncertainty is equivalent to consumption uncertainty. Therefore, lifetime utility is the expected value of lifetime income:

$$W = E \int_s^\infty U(\theta_s Y_s) e^{\delta t} dt = \frac{1}{\delta} e^{-\delta s} E[U(\theta_s Y_s)], \quad (4.2)$$

where δ is the rate of time preference.

By allowing $(\theta_s) = 1$ and $E[\theta_s - E(\theta_s)]^2 = \sigma_s^2$, and by applying a second-order Taylor series expansion, the expected utility of income,

$$E[U(\theta_s Y_s)] = E[U(Y_s)] + Y_s U'(Y_s) E(\theta_s - 1) + \frac{1}{2} Y_s^2 U''(Y_s) E(\theta_s - 1)^2, \quad (4.3)$$

simplifies to

$$E[U(\theta_s Y_s)] = U(Y_s) + \frac{1}{2} Y_s^2 U''(Y_s) \sigma_s^2. \quad (4.4)$$

By substituting (4.4) into (4.2), the consumer objective is to maximize lifetime income by choosing an amount of higher education, s ; that is,

$$\max_s W(s) = \frac{1}{\delta} e^{-\delta s} \left[U(Y_s) + \frac{1}{2} Y_s^2 U''(Y_s) \sigma_s^2 \right]. \quad (4.5)$$

Differentiating (4.5) with respect to s and rewriting yields the following first-order condition:

$$\varepsilon_s \left\{ \mu_s - \rho_s \sigma_s^2 \left(\mu_s + \gamma_s - \frac{1}{2} \delta \right) \right\} - \delta = 0, \quad (4.6)$$

with the marginal rate of return to schooling, $\mu_s = \frac{\partial Y_s}{\partial s} \frac{1}{Y_s} \geq 0$,

the relative risk gradient of risk to schooling, $\gamma_s = \frac{\partial \sigma_s}{\partial s} \frac{1}{\sigma_s}$,

relative risk aversion, $\rho_s = \frac{U''(Y_s)}{-U'(Y_s)} (Y_s)$,

and the income elasticity of utility, $\varepsilon_s = \frac{\partial U}{\partial Y_s} U(Y_s) > 0$.

By setting $\sigma_s^2 = 0$ and $\varepsilon_s = 1$, the first-order condition shown in (4.6) can be recognized as a generalization of the traditional human capital framework in which investment in education is predicted until the discount rate equals the marginal rate of return (Hartog & Diaz-Serrano, 2007, p. 7). From (4.6), individual attitudes towards risk, ρ_s , clearly have an impact on the optimal amount of education. The effects of changes in risk (γ_s) and changes in the returns to education (μ_s) depend on consumer attitudes towards risk. If risk increases with education, risk aversion (tolerance) is expected to act as a deterrent (incentive) to human capital investment; however, if risk decreases with education, the effect is reversed (Hartog & Diaz-Serrano, 2007, p. 7).

For the purposes of this study, an additional assumption is made. In Hartog and Diaz-Serrano's model, the consumer objective is to maximize lifetime income which is a function of the amount of realized schooling. However, when making the decision to pursue higher education, consumers do not know with certainty the amount that they will be able to successfully complete if undertaken (i.e., the amount realized). Given the sheepskin effects, intuition suggests that the likelihood of completion would affect a rational consumer's expectations of the returns to education. Therefore, assume that, in addition to the framework above, the likelihood of enrolling is assumed to be at least in part based on the subjective probability of completion. This assumption will be tested in the empirical analysis.

4.2 Empirical Model

Based on this theoretical framework, the following empirical model is derived. Let S^* , the amount of higher education chosen, be a latent variable.

$$S^* = \alpha_0 + \mathbf{X}'\beta + \mathbf{Z}'\gamma + \pi R + \varepsilon \quad (4.7)$$

where $\mathbf{X}'$ is a vector of individual-level control variables,

$\mathbf{Z}'$ is a vector of regional-level control variables,

R is relative risk tolerance, and

ε is an error term.

While the theoretical model specifies the marginal returns to education and the relative risk gradient, the current analysis implicitly controls for these variables through individual-level controls and regional control variables. Since the young consumers in this study are assumed to make the decision to enroll in higher education at approximately the same time, macroeconomic labor market opportunities should not vary substantially between individuals. Additionally, individual-level control variables include known predictors of earnings; therefore, this model indirectly controls for expected wages.

Given the latent variable model in (4.7), the observed consumer choice is the decision to enroll in higher education, S . If $S^* > 0$, then the consumer chooses to enroll in higher education ($S = 1$); if $S^* \leq 0$ then the consumer choose to not enroll in higher education ($S = 0$). That is,

$$S = \begin{cases} 1, & \text{if } S^* > 0 \\ 0, & \text{otherwise} \end{cases} \quad (4.8)$$

Therefore, the empirical model will make use of a binary response model and model the probability that $S = 1$ based on the set of independent variables in (4.7). The empirical model is discussed in greater detail in Section 5.4.1.

4.3 Hypotheses

Based on the theoretical framework and empirical model, the following research hypothesis is formed:

RH: Consumer risk preferences have an influence on the decision to enroll in a four-year institution of higher education.

Empirically, the research hypothesis will be tested by evaluating the null hypothesis, the coefficient of risk tolerance is zero ($H_0: \pi = 0$), against the alternative, that the coefficient of risk tolerance is not zero ($H_1: \pi \neq 0$). Rejecting H_0 gives supporting evidence to the hypothesis (i.e., risk preferences have an influence on enrollment decisions) while failing to reject the null indicates no evidence of the hypothesized relationship (i.e., risk preferences do not influence enrollment decisions). The non-directional hypothesis is justified by the ambiguous relationship between risk and higher education investments. That is, the literature shows that there is not a clear positive or negative relationship between risk and education (Hartog & Diaz-Serrano, 2007, p. 10-11).

Chapter 5: Research Design and Method

This chapter outlines the research design and methodology employed to answer the research question and is organized as follows. The dataset and sample selected for analysis are described in Section 5.1 and Section 5.2, respectively. Section 5.3 discusses the measurement of the dependent variable and independent variables. Section 5.4 covers the primary research methodology used to address the research question. The chapter concludes with Section 5.5, which describes several sensitivity analyses used to examine the robustness of the findings to the empirical specifications.

5.1 Data

A number of datasets have been used to study higher education decisions and outcomes. The National Center for Education Statistics (NCES) collects several large, nationally representative datasets that are commonly used among education researchers, including the Baccalaureate and Beyond Longitudinal Study, Beginning Postsecondary Students, High School and Beyond, and the National Postsecondary Student Aid Study (National Center for Education Statistics, 2013b). Additionally, the Bureau of Labor Statistics collects the National Longitudinal Survey of Youth (NLSY). There are currently two cohorts of the NLSY: the 1979 and the 1997 cohorts. The BLS also

collected a special dataset surveying the children of the 1979 cohort – the NLSY 1979 Children and Young Adults.

For the current study, the NLSY97 was chosen as the most appropriate dataset to address the research question for several reasons. As a longitudinal dataset, the NLSY97 allows researchers to follow young consumers over time, which includes periods before, during, and after higher education experiences. Because this study is concerned with the education decision making process, the ability to see the student's behavior during several points in time was important. The NLSY97 also collects a rich variety of information, including education, labor market participation, family background, health, and achievement tests (Bureau of Labor Statistics, 2013c). This variety is important because the theoretical model and literature suggest that many of these factors play an important role in education decisions. Furthermore, this dataset provides the greatest flexibility in using proxy variables to account for individual unobserved characteristics. Because the NLSY97 includes students who choose not to invest in higher education, this survey avoids many of the sample selection issues compared to the NCES datasets, which generally survey higher education students only.

As a nationally representative panel study of American youth, the NLSY97 was designed to provide information regarding the transition from school to work and adulthood (Bureau of Labor Statistics, 2013a). The first round of the NLSY97 was conducted in 1997, consisting of approximately 8,984 youth born between 1980 and 1984. The NLSY97 also includes an oversample of Hispanics and Blacks (Bureau of Labor Statistics, 2013c) which adds to the richness of the data. There have been 15 rounds of survey collection, with the most recent being released in August 2013. This

dataset is publically available at no cost and researchers may also request access to the NLSY97 Geocode Data which contains geographic information (e.g., state, MSA, etc.) and the ability to match institutes of higher education attended by respondents with the IPEDS data.

5.2 Sample

The NLSY97 was restricted in several ways to create a sample for this analysis. Because only students who graduate from high school or who have a GED are eligible for admission to higher education, the sample was restricted to include only high school graduates or those who completed a GED by the age of 20. The age restriction was added so that the focus is traditional undergraduates. This restriction was also similar in scope to previous research examining higher education enrollment decisions (e.g., Lovenheim & Reynolds, 2011). Furthermore, this definition is not overly restrictive since nearly two-thirds of high school graduates in the U.S. enroll in higher education in the fall term immediately following high school completion (National Center for Education Statistics, 2013a). Of the 8,984 participants in the NLSY97, 7,050 respondents were eligible to attend higher education in the relevant time frame. That is, 7,050 respondents had completed high school or a GED by the age of 20.

The sample was also restricted to include only respondents who were at least 15 years old as of December 31, 1996. After applying this restriction, the sample size was reduced to 2,781 participants. The reason for this restriction is that a key control variable (the subjective probability of college completion) was asked only of respondents who were at least 15 by the end of 1996. Given the importance of this variable, the sample

was restricted to those who were asked this survey question. Limiting the sample to this age range also helped ensure that survey questions that were asked only in Round 1 were asked when the respondents were approximately close to high school graduation, the point at which the decision to enroll in higher education is typically made (National Center for Education Statistics, 2013a).

For the primary analysis, the sample was also restricted to complete cases only; i.e., cases with missing data were dropped from the analysis. This reduced the number of cases from 2,781 to 1,041. This restriction may introduce nonresponse bias if the respondents who were not selected for analysis (the incomplete cases) were systematically different than those who were selected for the analysis (the complete cases) (Greene, 2012, p. 95). Subsequently, this study used several robustness checks to examine the sensitivity of results to possible nonresponse bias. Details regarding missing data and sensitivity analyses are discussed in Section 5.5.

Lastly, because the NLSY97 includes both a complex survey design and an oversample of minority groups, weighting is an important issue to consider for researchers making generalizations to the population. The NLS staff have created weights for researchers working with either individual survey years or multiple survey years that correct for both the complex design of the survey and the oversample of minorities (Bureau of Labor Statistics, n.d.-d). If multiple years of the NLSY97 are used, the NLS staff recommends creating a custom weight from the NLS website. At the website, researchers can select the survey years from which data are pulled and have a custom weight created for use in analysis. Data for the current study was drawn from 1997-2001,

2010, and 2011, therefore, a custom weight was created by selecting the corresponding survey years from the NLS site (i.e., 1997, 1998, 2000, 2001, 2010, and 2011).

5.3 Measurement of Variables

This section specifies the operationalization of the dependent variable (5.3.1) and the independent variables (5.3.2) used in the current study. One methodological challenge in studying the decision making process of consumers is that, in reality, there is not an exact decision point in which higher education choices are made once and for all. Since consumers have the ability to choose to delay higher education entry into the foreseeable future, simplifying assumptions must be made in order to assess the higher education choice in a meaningful way. The approximate time in which an initial decision is made is a critical input because time-variant explanatory variables need to be measured at (or close to) the time of the decision. Furthermore, a key variable of interest, risk aversion, is likely time-variant (Ahn, 2010; Light & Ahn, 2010; Sahm, 2012). The timing issue is complicated further by the fact that key variables were only asked in certain rounds (often Round 1) when students were different ages. Therefore, whenever possible, the independent variables were measured in Round 1, when students are either 15 or 16 years old. Since it is possible that this timing issue may affect the findings, a sensitivity analysis was used to examine the robustness of findings with respect to age (see Section 5.5). Unless otherwise noted, details regarding survey questions and coding of responses can be found at the NLS Investigator website (Bureau of Labor Statistics, n.d.-c).

5.3.1 Dependent Variable

Since the amount of higher education that the consumer has chosen to pursue at the initial enrollment decision is not observed (i.e., actual attainment and planned attainment may not be the same), this empirical analysis will focus on the decision to enroll in higher education. Therefore, the dependent variable in this study is a binary variable distinguishing consumers who enrolled in either a two-year or four-year school (coded 1) from those who did not enroll in either school by the age of 20 (coded 0). Although careful distinctions between two-year and four-year attendees are often made in applied research, given the latent variable model described in Section 4.2, the relevant comparison for the current study was between those who chose to attend a higher education institution and those who chose not to attend an institution.

5.3.2 Independent Variables

As indicated in the empirical model, two categories of control variables entered the model as predictors in the decision to enroll in higher education: individual-level controls and regional controls. The primary independent variable that was used to test the research hypothesis was a measure of risk tolerance. Each set of control variables and the key independent variable of interest are described in the discussion that follows. Table 5.1 summarizes each variable.

Table 5.1. Variable Names and Coding

Full Variable Name	Abbreviation	Description
DV: Higher Enrollment by Age 20	Enroll	= 1 if respondent enrolled in higher education by age 20; 0 otherwise.
INDIVIDUAL-LEVEL CONTROLS		
High School GPA	-	GPA on a four point scale, measured as a continuous variable.
<i>Family Structure</i>		
Both Biological Parents	-	= 1 if respondent lived with both biological parents; 0 otherwise.
Two Parents, One Biological	-	= 1 if respondent lived with two parents and one was a biological parent; 0 otherwise.
Biological Mother Only	-	= 1 if respondent lived with biological mother only; 0 otherwise.
Biological Father Only	-	= 1 if respondent lived with biological father only; 0 otherwise.
Other Parental Figures	-	= 1 if respondent lived with parent figures who were of some other relation; 0 otherwise.
<i>Parental Education</i>		
Less Than 12th Grade	-	= 1 if highest education level of respondent's parent(s) was less than a high school diploma; 0 otherwise.
12 th Grade	-	= 1 if highest education level of respondent's parent(s) was a high school diploma or equivalent; 0 otherwise.
Some College	-	= 1 if highest education level of respondent's parent(s) was less than four years of college; 0 otherwise.
Four Years of College	-	= 1 if highest education level of respondent's parent(s) was four years of college; 0 otherwise.
More than Four Years	-	= 1 if highest education level of respondent's parent(s) was more than four years of college; 0 otherwise.
Family Net Worth	-	Difference between assets and liabilities, measured as a continuous variable in \$10,000 increments.
Family Income	-	Total family income for 1996, measured as a continuous variable in \$10,000 increments.
Probability of College Degree by 30 According to Parent	Probability of Degree – Parent	Parent estimate of the probability that the respondent would have a college degree by the age of 30, measured as a continuous

Table 5.1: *Continued*

Full Variable Name	Abbreviation	Description
		variable in increments of 10%.
<i>Sex</i>		
Female	-	= 1 if respondent's reported sex was female; 0 if respondent's reported sex was male.
<i>Race/Ethnicity</i>		
White	-	= 1 if respondent's reported race/ethnicity was White only; 0 otherwise.
Black	-	= 1 if respondent's reported race/ethnicity was Black only; 0 otherwise.
Hispanic	-	= 1 if respondent's reported race/ethnicity was Hispanic; 0 otherwise.
Asian/Other	-	= 1 if respondent's reported race/ethnicity was non-White, Non-Black, and non-Hispanic; 0 otherwise.
Cognitive Ability Proxy: Armed Services Vocational Aptitude Battery	ASVAB	Percentile score created by the NLS staff, measured as a continuous variable ranging from 0% to 100%.
Time Preference Proxy: Ever Smoked	Ever Smoked	= 1 if respondent reported ever smoking; 0 otherwise.
Probability of College Degree by 30 According to Respondent	Probability of Degree	Self-estimate of the probability that the respondent would have a college degree by the age of 30, measured as a continuous variable in increments of 10%.
CENSUS REGION CONTROLS		
Midwest	-	= 1 if respondent resided in Midwest Census Region in 1997; 0 otherwise.
Northeast	-	= 1 if respondent resided in Northeast Census Region in 1997; 0 otherwise.
South	-	= 1 if respondent resided in South Census Region in 1997; 0 otherwise.
West	-	= 1 if respondent resided in West Census Region in 1997; 0 otherwise.
RISK PREFERENCES		
Age-Adjusted Coefficient of Relative Risk Tolerance	Relative Risk Tolerance	Measured as a continuous variable. Higher scores indicate that the respondent is more risk tolerant. See Appendix A for details.

Individual-level controls. A number of control variables were included in the empirical model to provide reliable estimates. Based on the theoretical framework and

the college enrollment, persistence, and returns to education literature, the following categories of individual-level control variables were created: academic ability, family background, expected wages, time preferences, and subjective probabilities.

Academic ability. There are a number of survey questions that can be used to measure academic ability among respondents. The NLSY97 collects detailed information from respondents regarding their educational history, including type and number of schools attended, courses taken, and achievement measures. Some of this data is collected from student transcripts provided that the student returned the consent for the school to share transcripts with the NLSY. For the current study, high school GPA was used to represent student academic ability. GPA was measured as a continuous variable taken from the student's high school transcript in 1999 that is standardized to a 4-point scale by the NLSY staff. Although the NLSY also collects SAT/ACT exam scores, these questions were not used to measure academic ability as a large number of respondents either did not take the exam or did not report their scores.

Family background. Family background included the respondent's relationship to parent figures with whom they reside, family net worth, family income, and parental education level. The relationship between the respondent and the parent figure or guardian with whom the respondent resides was collected in each round of the survey from until 2003. There were 10 response categories for this survey question in Round 1: (1) both biological parents, (2) two parents, biological mother, (3) two parents, biological father, (4) biological mother only, (5) biological father only, (6) adoptive parent(s), (7) foster parent(s), (8) no parents, grandparents, (9) no parents, other relatives, (10) anything else. Based on the distribution of the responses, this study condensed these categories

into five categories: (1) both biological parents, (2) step family, (3) biological mom, (4) biological dad, and (5) other. Dummy variables were created to represent each of the five categories.

Family net worth was computed by the NLS staff based on responses from the parent survey in 1997 regarding assets and liabilities (Bureau of Labor Statistics, n.d.-b). The parent survey is only collected in 1997 so in subsequent rounds of the survey, family net worth is based on responses from the youth. Family net worth is top-coded at \$600,000 in the NLSY to protect the privacy of survey participants. In this study family net worth was a scaled version of the computed net worth in the NLSY based on the 1997 parent survey; specifically the computed parental net worth was divided by 10,000 to provide easier interpretation of the estimated coefficients.

Family gross income was also computed by the NLS staff from responses to several NLSY questions that were collected in each round of the NLSY. For Round 1, family income was collected from the parent survey unless the youth was independent (i.e., he or she did not reside with parents). Therefore, a dummy variable was initially created to indicate that the gross income variable reflected the youth's income; youth who were considered independent were coded 1. However, after the sample restrictions were applied, fewer than 10 respondents were classified as independent in Round 1. Therefore, this variable was not included in the final version of Model 1. For the current study, the family income variable reported in the NLSY97 was divided by 10,000.

The NLSY97 also asks the students about the highest grade level completed by their residential parents and their biological parents. Responses ranged from 1st grade to 8th year of college or more. This study created five categories and used dummy variables

to represent the highest grade level completed by the residential parent(s): less than high school, high school, some college, four-year degree, and graduate degree. The fourth year of college was used as a cutoff point. Parents who completed four years of college were placed in the B.S. degree category; those who completed at least one year but less than four years of college were placed into the some college category, and those who completed more than four years were placed into the graduate degree category. Lastly, parents were asked for the probability that the respondent would have a college degree by the age of 30. A continuous variable was created based on responses to this question, which ranged from zero to 100 percent.

Expected wages. Expected wages are an important part of the human capital investment choice. Traditional earnings equations are a function of labor market experience, education, and ability. Because some students (but not all) start working in high school, experience is assumed to have a minimal impact on expected wages at the time of the decision to attend higher education. Education level is approximately the same for students because they are high school students. That leaves ability as a key explanatory variable for expected wages. In the NLSY97, a computer adaptive form of a military enlistment test, the Armed Services Vocational Aptitude Battery (ASVAB), was administered to respondents between 1997 and 1998. Each respondent was eligible to take the exam in Round 1 although some participants (approximately 21 percent) decided to not take the exam. The NLSY staff created a combined math and verbal percentile score for each participant who completed the ASVAB (Bureau of Labor Statistics, n.d.-a). This percentile score was used to proxy for ability.

The literature also suggests that sex and race need to be considered in expected wages due to earnings gaps between males and females and between Whites and non-Whites. A dummy variable distinguishing females (coded 1) from males (coded 0) was created. Dummy variables were also created to represent each respondent's racial/ethnic identity. Two questions from the NLSY97 were used to categorize respondents by racial/ethnic identity. The first question asked each respondent whether or not he or she was Hispanic. The second asked the respondent to identify his or her race from the following options: (1) White, (2) Black or African American, (3) American Indian, Eskimo, or Aleut, (4) Asian or Pacific Islander, (5) other. This study created dummy variables to distinguish between the following categories of respondent racial/ethnic identities: White, Black, Hispanic, and Asian/Other. Each of these variables was coded 1 if the respondent was in that racial/ethnic group and 0 otherwise. If a respondent reported being Hispanic, he or she was coded as Hispanic and all other categories were set equal to 0, although it is possible that the respondent also identified another race.

Time preferences. The theoretical model discussed in Section 4.1 posits that the subjective discount rate affects the optimal amount of education since the returns to education are delayed until after the investment and are spread over the working life. While there is not a direct measure of time preference available in the NLSY97, there are a number of possible proxies. Previous research has found that smokers have higher rates of time preference than nonsmokers, although beginning in the late 1990s, the decision to smoke also captures phenomena such as SES and self-control (Scharff & Viscusi, 2011). For the current study, whether or not the respondent had ever smoked was used as a proxy for individual time preferences. Although there are certainly limitations in using

this proxy (as with any proxy), smoking was chosen as the best-available proxy for time preferences. Therefore, a dummy variable was created to distinguish those who reported ever smoking (coded 1) from those who reported never smoking (coded 0).

Subjective probability of college completion. The NLSY97 asks participants a number of questions about expectations. One question asks the student for the likelihood that they will obtain a college degree; the question reads as follows:

“Now think ahead to when you turn 30 years old. What is the percent chance that you will have a four-year college degree by the time you turn 30?”

As previously mentioned, this question is asked in 1997 but only if the respondent was at least 15 years old as of December 31, 1996. This question was used to create a variable representing the student’s subjective probability of completing college and ranges from 0 to 100 percent.

Regional controls. There are several reasons why it is important to control for regional influences in the decision to attend higher education. Human capital theory posits that the cost of foregone earnings is an important consideration in the decision to acquire higher education. Therefore, the local economic environment and labor market conditions (e.g., unemployment rates) need to be controlled for when estimating the likelihood of college enrollment. Labor market opportunities and corresponding education requirements also vary by region. A crude control variable can be created from the NLSY97 which collects information regarding the respondents region of residence in each survey round. This study used dummy variables to represent the respondent’s location during Round 1: Northeast, Midwest, South, and West. Figure 5.1 shows the region on a map of the United States. Please note that the NLSY97 codebook refers to the

North Central region, but this region was renamed the Midwest census region beginning June 1984 (U.S. Census Bureau, n.d.).

Figure 5.1. Map of U.S. Census Regions
Source: U.S. Census Bureau, available at U.S Department of Transportation (n.d.)

Risk preferences. As shown in the theoretical framework, risk preferences are a critical variable in the decision to pursue risky human capital. Given the importance of this variable in the current analysis of enrollment decisions, a potentially serious complication arises when working with the NLSY97 due to the timing of the risk

questions. The risk preference questions were not asked in the initial survey years of the NLSY97 where most consumers make the decision to pursue higher education. Instead, the risk questions were not asked until 2010 in Round 14 (Bureau of Labor Statistics, 2013b) when most respondents were between 26 and 30. The risk questions were also asked in Round 15, but only of those who were not interviewed in Round 14.

One solution to the timing issue is to assume that risk preferences are time-invariant. For the current study, four alternate risk measures were available from the NLSY97 if risk preferences are assumed to be time invariant (discussed in section 5.5). This assumption should be made with caution since research has shown that risk preferences do seem to vary with time (Ahn, 2010; Light & Ahn, 2010; Sahm, 2012), although time-invariant heterogeneity in risk preferences accounts for over 70 percent of the variation in risk preferences (Sahm, 2012).

An alternative, and arguably, a more appropriate solution to the timing issue is to use an age-adjusted measure of risk preferences (e.g., Light & Ahn, 2010). Since the NLSY-79 asks an income gamble question in multiple rounds, Light and Ahn (2010) were able to estimate how risk preferences change over time. Light and Ahn's method allows researchers to compute risk preferences at any point in time for a given individual. For the primary analysis, the current study computed an approximate measure of age-adjusted relative risk tolerance following Light and Ahn's method. The first step in developing the age-adjusted measure was to use the responses to the income gamble questions in the NSLY97.

Income gamble questions. Respondents in the NSLY97 were given an income gamble question and a follow up income gamble question that can be used to infer (see

Appendix A) individual coefficients of relative risk tolerance. The initial income gamble question asked respondents the following question:

“Now I have another kind of question. Suppose you are the only income earner in the family, but that your current job is ending. You have to choose between two new jobs. The first job would guarantee your current family income for life. The second job is also guaranteed for life and possibly better paying, but the income is less certain. There is a 50-50 chance that the second job will double your current family income for life and a 50-50 chance that it will cut your current family income by a third for life. Which job would you take: the first job or the second job?”

Respondents who answered “first job” to the first question were asked this follow up question:

“Suppose the chances were 50-50 that the second job would double your current family income and 50-50 that it would only cut it by 20 percent. Would you take the first job or the second job?”

Respondents who answered “second job” to the first question were asked this follow up question:

“Suppose the chances were 50-50 that the second job would double your current family income and 50-50 that it would cut it in half. Would you take the first job or the second job?”

Based on responses to the initial and follow up income gamble questions, an ordinal measure of risk preferences can be created (Barsky, Juster, Kimball, & Shapiro, 1997; Light & Ahn, 2010). Respondents who chose the first job in both the first question and the follow up question were assigned to the first (lowest) risk tolerance category. Respondents who chose the first job in the first question and the second job in the follow up question were assigned to the second risk tolerance category. Respondents who chose the second job in the first question and the first job in the follow up question were assigned to the third risk tolerance category. Lastly, respondents who chose the second

job in both the first question and follow up question were assigned to the fourth (highest) risk tolerance category. Figure 5.2 summarizes the income gamble questions and corresponding risk tolerance categories.

Figure 5.2. Income Gamble Question Branching and Risk Categories

Age-adjusted risk tolerance. As previously mentioned, recent research suggests that risk preferences may vary with time – specifically, that people become more risk

averse (i.e., less risk tolerant) as they age (e.g., Light & Ahn, 2010; Sahm, 2012). Ideally, Light and Ahn's method would be replicated using the NLSY97 to develop an age-adjusted risk measure for the point in time at which the higher education enrollment choice was made. However, their method is currently only possible with the NLSY79 since that is the only NLSY cohort in which respondents have been asked the risk preference questions in multiple years. Respondents in the NLSY97 have only been asked the risk preference questions one time. Therefore, this study used the parameter estimates from Light and Ahn's NLSY79 study to approximate the coefficient of relative risk tolerance for consumers in the NLSY97 at the time they choose to enroll in higher education.

Light and Ahn (2010, p. 902) used the income gamble question that was asked in several rounds of the NLSY79 and estimated the coefficient of relative risk tolerance, ρ , for person i at time t , which depends on age, an individual-specific random effect (α_i), and a time-variant measurement error (u_{it}):

$$\log \rho_{it} = \beta \text{AGE}_{it} + \alpha_i + u_{it}. \quad (5.1)$$

Additionally, (5.1) was estimated separately for men and women based on the literature that suggests there are systematic differences between male and female preferences for risk. Light and Ahn's estimates are reported in Table A.2 in Appendix A. These estimates were used along with (5.2) to estimate the age-adjusted coefficient of relative risk tolerance for consumers in the NLSY97 at the time they choose to enroll in higher education (see Light & Ahn, 2010, p. 918):

$$E[\rho_{it} | c_{it} = j, \text{AGE}_{it}, \text{AGE}_{it}] = \exp\left(\hat{\alpha} + \hat{\beta} \text{AGE}_{it} + \frac{1}{2} \hat{\sigma}_{\alpha}^2\right) \frac{P(\log \underline{\rho}_j < \log \rho_{it} + \hat{\sigma}_{\alpha}^2 < \log \bar{\rho}_j)}{P(\log \underline{\rho}_j < \log \rho_{it} < \log \bar{\rho}_j)}, \quad (5.2)$$

where c_{it} = relative risk tolerance category (j) for person i at the interview date, time t ,

AGE_{it} = age of person i at the interview date, time t ,

$AGE_{i\tau}$ = age of person i at the decision date, time τ ,

$\underline{\rho}_j$ = lower bound for coefficient of relative risk tolerance for category j ,

$\bar{\rho}_j$ = upper bound for coefficient of relative risk tolerance for category j .

The decision date (time τ) was set to the first date of enrollment for those who enrolled in higher education and to either the date of high school graduation or GED completion for those who did not enroll in higher education. A discussion regarding the upper and lower bounds in (5.2) and further details regarding the calculation of the age-adjusted risk tolerance measure can be found in Appendix A.

This age-adjusted risk tolerance measure should be viewed cautiously for two reasons. For one, the income gamble question asked multiple years in the NLSY79 is worded slightly differently in the NLSY97. The lotteries and probabilities in the questions between the two cohorts are the same, but the question is framed slightly differently. The NLSY79 frames the choice as between the respondent's current job and a new job while the NLSY97 frames the choice as between two new jobs. Some respondents may simply want to avoid taking a new job or feel resistant to change which would bias the estimate of risk aversion. The change from a current job and new job to two new jobs was likely done to eliminate status quo bias (see Barsky et al., 1997). Furthermore, there is a timing issue since the NLSY79 respondents are approximately 10

years older when they are first asked the income gamble question compared to the age when the question is asked in the NLSY97.

Despite these issues, using Light and Ahn's estimates is currently the only available method for creating an age-adjusted measure of risk tolerance when working with the NLSY97. Therefore, this study acknowledged these limitations and used Light and Ahn's estimates as a rough approximation of risk preferences at the time of the decision to enroll in higher education. Several robustness checks, discussed in Section 5.5, were also conducted to check how sensitive the results were to the measurement of risk tolerance.

5.4 Research Method

This section describes the primary analysis used to evaluate the research question (Section 5.4.1) and addresses the implications of the literature implying gender differences in areas such as human capital accumulation and risk preferences (Section 5.4.1).

5.4.1 Model 1

Given the binary nature of the response variable, either the logit or probit model is well-suited to examine the research questions. The primary difference between the two models is the function that is used to limit the estimated response probabilities to values between zero and one; the logit uses the standard normal logistic function while the probit uses the normal cumulative distribution function (Wooldridge, 2009). The logit model was chosen over the probit given the relatively simple interpretation of model

coefficients as odds ratios (Allison, 2012) although both models would yield the same conclusions regarding the research hypothesis. Generally, the basic logistic regression model predicts the probability that the response variable (Y) is equal to one given a set of explanatory variables (X') using the logistic distribution (Wooldridge, 2009):

$$Prob(Y = 1|X') = \frac{\exp(X'\beta)}{1+\exp(X'\beta)} = \Lambda(X'\beta). \quad (5.3)$$

In order to estimate the empirical model derived in Chapter 4, logistic regression was used to model the likelihood of enrolling in an institution of higher education, $Prob(S = 1)$ given the set of previously described explanatory variables; that is,

$$Prob(S = 1|(\alpha_0 + \mathbf{X}'\beta + \mathbf{Z}'\gamma + \pi R)) = \frac{\exp(\alpha_0 + \mathbf{X}'\beta + \mathbf{Z}'\gamma + \pi R)}{1 + \exp(\alpha_0 + \mathbf{X}'\beta + \mathbf{Z}'\gamma + \pi R)}. \quad (5.4)$$

where α_0 is an intercept term,

$\mathbf{X}'$ is a vector of individual-level control variables,

$\mathbf{Z}'$ is a vector of regional-level control variables, and

R is the estimated relative risk tolerance.

By setting Λ equal to the cumulative logistic distribution function, (5.4) can be rewritten as

$$Prob(S = 1|(\alpha_0 + \mathbf{X}'\beta + \mathbf{Z}'\gamma + \pi R)) = \Lambda(\alpha_0 + \mathbf{X}'\beta + \mathbf{Z}'\gamma + \pi R). \quad (5.5)$$

The complete regression equation is shown in (5.6).

$$\begin{aligned} Enroll = & \alpha_0 + \beta_1(hsGPA) + \beta_1(Step) + \beta_2(BioMom) + \beta_3(BioDad) + \\ & \beta_4(OtherPar) + \beta_5(hsDrop) + \beta_6(SC) + \beta_7(BS) + \beta_8(Grad) + \beta_9(FamNW) + \\ & \beta_{10}(FamInc) + \beta_{11}(ParExp) + \beta_{12}(Female) + \beta_{13}(Black) + \beta_{14}(Hispanic) + \\ & \beta_{15}(AsianOther) + \beta_{16}(ASVAB) + \beta_{17}(Smoke) + \beta_{18}(PrGrad) + \gamma_1(NE) + \\ & \gamma_2(South) + \gamma_3(West) + \pi_1(RRT) . \end{aligned} \quad (5.6)$$

The logistic regression procedure in SAS 9.3 was used to generate maximum likelihood estimates of (5.6). The output from this procedure includes the estimate of the coefficient, the standard error, a chi-square test and corresponding p-value for each coefficient, and an estimate and confidence interval of the odds ratios. The estimated coefficients and corresponding chi-square test for the parameters in (5.6) were then used to draw conclusions regarding the research hypothesis discussed in Section 4.3.

To aid in the interpretation of the logistic regression, average marginal effects (also known as average partial effects) were computed using the margins command in STATA 11. Average marginal effects are an estimate of the *ceteris paribus* influence of a predictor variable on the explanatory variable and are computed as the summation of the marginal effect for each individual in the sample averaged over the sample (Wooldridge, 2010). See Appendix C for details regarding the calculation of average marginal effects.

5.4.2 Gender Differences

Since the literature has shown that gender plays a very important role in higher education decision patterns (for a review, see DiPrete & Buchmann, 2013) as well as being associated with differences in risk preferences and the returns to education, this analysis also examined whether or not it was necessary to estimate (5.6) separately for men and women. A likelihood ratio test was used to test the appropriateness of separating men and women for the analysis described in Section 5.4.1.

The likelihood ratio test can be used to test the null hypothesis that a restricted model is more appropriate than an unrestricted model. This test is similar to the Chow test in OLS regression settings that can be used to test for structural breaks (Greene,

2012, p. 706). The null hypothesis of the likelihood ratio test posits that q parameters in the unrestricted model should be restricted to zero (i.e., the restricted model is a better fit). Under the null hypothesis, the likelihood ratio test statistic has an approximate chi-square distribution with q degrees of freedom (Wooldridge, 2009). The likelihood ratio is computed by taking twice the negative difference of the log-likelihood functions of the restricted model ($\ln \hat{L}_R$) and the unrestricted model ($\ln \hat{L}_{UR}$) (Greene, 2012, pp. 703-706), that is

$$LR = -2[\ln \hat{L}_R - \ln \hat{L}_{UR}] \quad (5.5)$$

To test the appropriateness of separating men and women, the restricted model (Model 1), can be thought of as a pooled model since it includes both men and women and uses a dummy variable to control for sex. The unrestricted model allowed all of the coefficients in Model 1 to vary by gender. This was accomplished by running the logistic regression models (Model 1 less the dummy for sex) separately for men and women. This method is equivalent to running one model in which the gender dummy variable is interacted with each of the other parameters in the model. Hence, the likelihood ratio test statistic can be used to test whether the coefficients need to be estimated separately for men and women.

The restricted model (Model 1) had 23 parameters while each of the separate logistic regressions included 22 parameters (Model 1 less the dummy for sex), for a total of 44 parameters in the unrestricted model. Therefore, this approach tested 21 exclusion restrictions ($q = 21$), and the chi-square test statistic was

$$LR = -2[\ln \hat{L}_P - (\ln \hat{L}_M - \ln \hat{L}_W)] = \chi_q^2 \quad (5.6)$$

where $ln\hat{L}_p$ =log-likelihood function of pooled model,

$ln\hat{L}_M$ =log-likelihood function for men

$ln\hat{L}_W$ =log-likelihood function for women.

5.5 Sensitivity Analysis

A sensitivity analysis was conducted to examine the robustness of the findings to specifications in the empirical analysis. Specifically, there were three sources of concern that may introduce noise or bias into the analysis: (1) missing data (Section 5.5.1), (2) the measurement of risk tolerance (Section 5.5.2), and (3) the timing of variable measurement (Section 5.5.3).

5.5.1 Missing Data

As with many large surveys, missing data is an important consideration when working with the NLSY97. In practice, most researchers needing to deal with missing data simply delete cases that have missing observations and analyze complete cases only (Cameron & Trivedi, 2005, p. 925). However, missing data can cause serious problems in multivariate analysis because it may introduce bias into the estimates depending upon the missing data mechanism (i.e., the assumption regarding why the data is missing). Missing data mechanisms fall into three categories: (1) missing completely at random (MCAR), (2) missing at random (MAR), and (3) missing not at random (MNAR) (Greene, 2012, pp. 94-97).

The best-case scenario is in the case of MCAR, which assumes that observations are missing on a random basis and the resulting subsample of complete observations is a

random sample of the full sample (Cameron & Trivedi, 2005, p. 927). Since the resulting sample is a random sample of the full sample, cases with missing data can be eliminated without serious consequences. In the case of MAR, the likelihood of an observation being missing does not depend on its own value, but may depend on the values of other observable variables (Cameron & Trivedi, 2005, p. 926). If data is assumed to be MAR, there are no issues with nonresponse bias but the estimates may not be efficient (Cameron & Trivedi, 2005, p. 926). However, if MAR is violated, the value of the missing observation affects the likelihood of missingness and the missing data mechanism is said to be MNAR.

MNAR introduces nonresponse bias into the estimates because the complete cases are systematically different than the incomplete cases – i.e., there is a selection problem and the sample is not a random sample of the population of interest. Since there is not a direct test of assumptions regarding missing data, researchers must think carefully through missing data issues and use sensitivity analyses to test assumptions (Cameron & Trivedi, 2005, p. 926). Once the missing data mechanisms have been considered, there are various methods of imputing data or using sample selection models to account for missing data problems (see Cameron & Trivedi, 2005, Chapter 27; Greene, 2012, pp. 94-97).

For the current study, four variables are of particular concern. Cognitive ability, high school GPA, family net worth, and family income have a high number of missing observations and are likely MNAR; i.e., the value of each of these variables likely influences whether or not the observation is missing. For example, both high-income (net worth) and low-income (net worth) households may be less likely to report their income

(net worth); the value of the variable influences whether or not the observation is missing. Similarly, if the NLS staff was unable to gather the respondent's transcript information regarding GPA, it was likely due to inability to track the student's academic history or locate the school the respondent reported attending. These students may be lower-performing students on average. Lastly, since respondents chose whether or not to take the military aptitude test, it is plausible that lower-ability students declined the opportunity to take the exam.

Two sensitivity analyses were conducted to understand how missing data might affect the findings. The first analysis involved separating the sample into incomplete and complete case samples. A descriptive analysis comparing the two samples was conducted to determine whether the incomplete cases were systematically different from the complete cases. The second analysis involved an ad hoc imputation method to determine whether or not the inclusion of missing cases would alter the findings regarding the key variable of interest. While the primary analysis for the current study (Model 1) analyzed complete cases only, the results may be biased if the missing data is MNAR. Since a primary concern about data that is MNAR is that the sample is no longer a random sample, an ad hoc imputation method was used as a robustness check to see whether the results of the study were sensitive to restricting the sample to complete cases only.

Specifically, Model 1 was run a second time with imputed missing data replacing missing values (Model 1.IMP) for the variables that were of particular concern: ability, high school GPA, family net worth, and family income. For each of these variables, mean imputation was used to replace missing values and dummy variables were created to indicate whether or not the corresponding continuous variable was imputed. For example,

if GPA was missing, the missing value was replaced with the sample mean for high school GPA and a dummy variable for missing GPA was coded 1. Table 5.2 summarizes the variables created for Model 1.IMP. If those with missing information for variables of interest were systematically different than those without missing information, the key estimated coefficients for Model 1.IMP may differ from the findings of Model 1 and the dummy variables may capture some of the systematic differences. While mean imputation is generally not recommended for regression analyses (Cameron & Trivedi, 2005, p. 929), this approach will simply serve as a robustness check to determine whether or not the missing data problem is important enough to alter the evaluation of the research hypothesis.

Table 5.2. Variable Names and Coding for Model 1.IMP

Full Variable Name	Abbreviation	Description
SENSITIVITY ANALYSIS (Model 1.IMP)		
High School GPA – Imputed	GPA: Imputed	= high school GPA if observed; mean of high school GPA if missing.
High School GPA Missing	GPA: Missing	= 1 if high school GPA was missing; 0 otherwise.
Family Net Worth - Imputed	NW: Imputed	= family net worth if observed; mean of family net worth if missing.
Family Net Worth Missing	NW: Missing	= 1 if family net worth was missing; 0 otherwise.
Family Income - Imputed	INC: Imputed	= family income if observed; mean of family income if missing.
Family Income Missing	INC: Missing	= 1 if family income was missing; 0 otherwise.
Cognitive Ability Proxy: Armed Services Vocational Aptitude Battery - Imputed	ASVAB: Imputed	= ASVAB score if observed; mean of ASVAB score if missing.
Cognitive Ability Proxy – Armed Services Vocational Aptitude Battery Missing	ASVAB: Missing	= 1 if ASVAB score was missing; 0 otherwise.

5.5.2 Measurement of Risk Tolerance

Since the age-adjusted relative risk tolerance measure relied on parameter estimates of an older population and on a slightly different income gamble question, there may be concerns about the empirical findings in Model 1. Therefore, a sensitivity analysis was conducted to see whether alternative measures of time-invariant risk preferences available in the NLSY97 would affect the conclusions regarding the research hypothesis. Model 2 included dummy variables created from the income gamble

categories described in Section 5.3.2. This approach is justifiable if risk preferences are assumed to be time-invariant.

In addition to the income gamble question, a general willingness to take risks question and several context-specific risk questions were asked. The general willingness to take risks asked the following question:

“Are you generally a person who is fully prepared to take risks or do you try to avoid taking risks? Rate yourself from 0 to 10, where 0 means ‘unwilling to take any risks’ and 10 means ‘fully prepared to take risks.’”

The general willingness to take risks question was then applied to several specific contexts and asked in a similar way:

“How would you rate your willingness to take risks in the following areas? [while driving, in financial matters, in respondent’s occupation, with respondent’s health, in respondent’s faith in other people, in romantic relationships, in making major life changes, and in placing bets]. For each situation, rate your willingness from 0 to 10...”

These questions were used to develop three continuous measures of risk tolerance, ranging from 0 to 10. The three contexts chosen to be most closely related to the decision to enroll in higher education were the respondent’s willingness to take risk in general (*GenRisk*), in finances (*FinRisk*) and in making major life changes (*ChgRisk*).

After creating the alternate measures of risk tolerance, Model 1 was rerun by replacing the age-adjusted relative risk tolerance measure with the alternate measures. This analysis compared estimates of the effect of risk tolerance by running Model 1 with four alternate risk measures: four dummy variables for each of the risk categories from the income gamble question (Model 2), the general willingness to take risks scale (Model 3), the financial willingness to take risks scale (Model 4), and the major change willingness to take risks scale (Model 5).

Table 5.3. Variable Names and Coding for Models 2 – 5

Full Variable Name	Abbreviation	Description
SENSITIVITY ANALYSIS (Models 2-5)		
<i>Risk Tolerance Categories</i>		
Relative Risk Tolerance – 1	RRT1	= 1 if respondent chose the first job in both income gamble questions; 0 otherwise. RRT1 indicates low risk tolerance.
Relative Risk Tolerance – 2	RRT2	= 1 if respondent chose the first job in the first income gamble question and the second job in the second income gamble question; 0 otherwise.
Relative Risk Tolerance – 3	RRT3	= 1 if respondent chose the second job in the first income gamble questions and the first job in the second income gamble question; 0 otherwise.
Relative Risk Tolerance – 4	RRT4	= 1 if respondent chose the second job in both income gamble questions; 0 otherwise. RRT4 indicates high risk tolerance.
Willingness to Take Risks in General	GenRisk	Measured as a continuous variable on a 0 – 10 scale, with higher (lower) levels indicating greater (less) willingness to take risks.
Willingness to Take Risks in Finances	FinRisk	Measured as a continuous variable on a 0 – 10 scale, with higher (lower) levels indicating greater (less) willingness to take risks.
Willingness to Take Risks in Major Life Changes	ChgRisk	Measured as a continuous variable on a 0 – 10 scale, with higher (lower) levels indicating greater (less) willingness to take risks.

5.5.3 Timing of Variable Measurement

Lastly, the timing of survey questions is not uniform among the respondents because respondents ranged from ages 15 to 16 at the end of 1996. There may be concerns that the predictors affect the dependent variable differently depending on the age of the respondent in 1997 because the measurement is farther from the actual

enrollment decision. Therefore, a likelihood ratio test (previously described in Section 5.4.2) was used to test for a structural break by running Model 1 separately for respondents who were less than 16 years old (young sample – Model 1.Y) and respondents who were 16 or older (old sample – Model 1.O).

To test the appropriateness of separating the sample in young and old, the restricted model, Model 1, can be thought of as a pooled model since it included both age groups. By running the logistic regression models (Model 1) separately for the two age groups, the unrestricted model allowed all of the coefficients in Model 1 to vary by age. The restricted model and each of the separate logits had 22 parameters; thus the unrestricted model had a total of 44 parameters. Therefore, this approach tested 22 exclusion restrictions ($q = 22$), and the chi-square test statistic was

$$LR = -2[\ln\hat{L}_P - (\ln\hat{L}_Y - \ln\hat{L}_O)] = \chi_q^2, \quad (5.8)$$

where $\ln\hat{L}_P$ =log-likelihood function of the pooled model,

$\ln\hat{L}_Y$ =log-likelihood function for the younger sample

$\ln\hat{L}_O$ =log-likelihood function for the older sample.

Chapter 6: Results

This chapter presents the findings of the current study. Section 6.1 presents the descriptive statistics of the sample. Section 6.2 presents the results of the multivariate analyses. Section 6.3 presents the results of the sensitivity analyses conducted to examine the robustness of the findings from Section 6.2.

6.1 Descriptive Results

This section presents the estimated population proportions and means from the sample used in the multivariate analysis. All proportions and means reported in Table 6.1 were weighted using the custom weight from the NLSY and a total of 1,041 cases were analyzed after applying the sample restrictions and dropping incomplete cases (see Section 5.2). Table 6.1 also presents the proportions and means by the dependent variable, i.e., higher education enrollment by age 20.

Three types of hypothesis tests were used to determine whether the sample proportions and means were statistically different than expected. For continuous variables, independent samples *t*-tests (see Field & Miles, 2010, pp. 280-283) were used to test whether the mean among those who did not enroll (Enroll=0) were significantly different than the mean among those who did enroll (Enroll=1). Continuous variables were also split into four categories in order to further examine descriptive differences.

Categories were generally based on quartiles of the continuous variable unless the distribution suggested an alternative scheme was necessary (e.g., subjective probabilities). For all categorical variables, a chi-square test of homogeneity (see DeGroot & Schervish, 2012, pp. 648-649) was used to test whether the observed variable proportions by enrollment were significantly different than expected under the null hypothesis, which was the full sample proportion for each variable. Lastly, for the last column in Table 6.1, a binomial proportion test was used to test whether the row percentage of $\text{Enroll}=1$ for each variable differed significantly from the overall sample proportion of $\text{Enroll}=1$. See Appendix B for details regarding each of these tests.

Approximately 65 percent of young adults had enrolled in higher education by the age of 20. The mean GPA in high school was 2.88. Among the four levels of GPA, there are large, significant differences in the proportion of young adults who enroll compared to the overall sample proportion (65 percent). For example, for young adults with GPAs below 2.53, only 39 percent enrolled in higher education; for young adults with GPAs over 3.35, nearly 95 percent enrolled in higher education.

Regarding family structure, the majority of young adults lived with either their biological parents (57 percent) or biological mother (23 percent). Significantly higher proportions of young adults who enrolled in higher education lived with both biological parents compared to those who did not enroll. Larger proportions of those who did not enroll lived with one parent or other parent figures compared to those who did enroll. For example, young adults who lived with the biological father only enrolled at significantly lower rates (42 percent) than the sample proportion (65 percent).

For nearly 40 percent of young adults, the highest education level of parents was a high school diploma (29 percent) or less (11 percent). Approximately 29 percent of respondents' parents spent less than four years in college; 17 percent of parents spent four years in college, and the remaining 14 percent spent more than four years in college. Compared to the enrollment proportion for the full sample, young adults from less educated households (either high school diploma or less) enrolled in significantly lower proportions (42 and 46 percent, respectively) while those from more educated households, ranging from at least one year of college to eight years or more, had significantly higher enrollment rates (between 72 percent and 87 percent) than the enrollment rate for the full sample.

The mean family net worth reported in 1997 was \$134,412 and the mean family income was \$56,530. For both family net worth and family income, young adults in the highest (lowest) quartile had significantly higher (lower) enrollment rates than the enrollment rate for the full sample. Parents generally had high expectations that young adults would complete a college degree by the age of 30 – the mean probability was 75 percent.

The sample had a slightly higher proportion of females (51 percent) compared to males (49 percent). Enrollment rates were significantly lower for males (59 percent) than the average enrollment rate and significantly higher for females (71 percent). In terms of race/ethnicity, 74 percent of young adults were White, 11 percent were Black, nearly 10 percent were Hispanic, and 4 percent were Asian or some other race/ethnicity. Hispanics had significantly lower enrollment rates (48 percent) while Asians and others (85 percent) had significantly higher enrollment rates than the average.

In terms of cognitive ability, as measured by the ASVAB score, the mean percentile was 56 percent. Large, significant differences in enrollment rates were observed between the sample enrollment rate and different levels of cognitive ability. For example, among the lowest quartile of cognitive ability, only 35 percent enrolled; in the highest quartile, nearly 90 percent enrolled. Approximately 52 percent of young adults reported ever smoking, which is interpreted as having relatively higher rates of time preferences. Despite this variable being a relatively crude proxy for time preference, those who had never smoked had significantly higher enrollment rates (72 percent) and those who had smoked had significantly lower enrollment rates (59 percent) than the sample average. Young adults tended to have high expectations of college completion by the age of 30: the mean was 79 percent.

Over half of young adults were located in the Midwest (34 percent) or South (31 percent) census regions; 20 percent were in the West and the remaining 15 percent were in the Northeast. There were not significant differences between the various regions and the sample enrollment rate.

The mean age-adjusted coefficient of relative risk tolerance (RRT) was 0.97. The mean level of relative risk tolerance was significantly lower among those who did not enroll compared to those who did enroll. Table 6.2 presents the mean age-adjusted RRT by each quartile of risk tolerance. For comparison purposes, Table 6.3 presents the estimated coefficient of relative risk tolerance from Light and Ahn's (2010) study. Although the means were comparable, the range of the estimated risk tolerance for the current study is substantially smaller than in Light and Ahn's study. The reason for this is that Light and Ahn were able to estimate individual-specific random effects because they

had multiple survey years; the current study had to use the average individual-specific random effect for every individual, which greatly reduced the variation. The current study also had a much lower variation due to the comparatively small sample size.

Table 6.1. Descriptive Statistics of Complete Cases

Variable	Sample Proportion (Mean)			Significance Test ^b			Row % Enroll=1 ^c
	Column % Full Sample N=1,041 ^a	Column % Enroll=0	Column % Enroll=1	χ^2 (t) Test Statistic	df	p- value	
DV: Enroll	64.87%	0%	100%	-			-
INDIVIDUAL-LEVEL CONTROLS							
High School GPA	2.88	2.50	3.09	-14.97 ^t	1039	<.01	-
2.53 and under	24.84%	43.01%	15.00%	74.98	1	<.01	39.17%**
2.54 – 2.94	24.74%	32.64%	20.47%	14.20	1	<.01	53.67%**
2.95 – 3.34	24.92%	20.38%	27.38%	4.67	1	0.03	71.28%*
3.35 and over	25.50%	3.97%	37.15%	102.45	1	<.01	94.53%**
<i>Family Structure</i>	-	-	-	41.84	4	<.01	
Both Biological Parents	56.53%	44.21%	63.20%	15.12	1	<.01	72.53%**
Two Parents, One Biological	14.13%	15.56%	13.36%	0.82	1	0.37	61.32%
Biological Mother Only	23.32%	30.34%	19.52%	11.92	1	<.01	54.29%**
Biological Father Only	3.23%	5.36%	2.07%	7.94	1	<.01	41.68%**
Other Parental Figures	2.79%	4.52%	1.86%	6.94	1	<.01	43.12%*
<i>Parental Education</i>	-	-	-	128.78		<.01	
Less Than 12th Grade	10.54%	17.55%	6.74%	26.31	1	<.01	41.50%**
12 th Grade	29.37%	44.87%	20.97%	46.12	1	<.01	46.33%**
Some College	28.83%	22.34%	32.34%	8.23	1	<.01	72.78%**
Four Years of College	17.03%	9.89%	20.89%	16.88	1	<.01	79.61%**
More than Four Years	14.24%	5.36%	19.05%	31.25	1	<.01	86.79%**
Household Net Worth (\$)	134,412	72,204	168,094	-7.85 ^t	1039	<.01	
20,299 and below	24.93%	37.20%	18.29%	34.06	1	<.01	47.58%**
20,300 – 75,099	25.02%	31.98%	21.25%	10.91	1	<.01	55.10%**
75,100 – 184,499	24.93%	20.71%	27.21%	4.02	1	.04	70.82%*
184,500 and over	25.13%	10.11%	33.26%	50.58	1	<.01	85.87%**
Household Income (\$)	56,530	40,353	65,289	-9.54 ^t	1039	<.01	
29,199 and below	24.89%	38.27%	17.65%	40.53	1	<.01	46.00%**
29,200 – 47,499	25.11%	27.69%	23.71%	1.49	1	0.22	61.27%
47,500 – 70,699	24.94%	20.69%	27.24%	4.08	1	0.04	70.86%*
70,700 and over	25.06%	13.35%	31.40%	30.83	1	<.01	81.29%**
Probability of Degree – Parent	75.10%	58.66%	84.01%	-14.87 ^t	1039	<.01	
0% – 49%	10.92%	24.25%	3.70%	91.82	1	<.01	21.96%**
50%	37.85%	46.46%	33.19%	11.04	1	<.01	56.88%**
51% – 99%	12.57%	7.53%	15.39%	12.22	1	<.01	79.47%**
100%	38.67%	21.94%	47.72%	40.78	1	<.01	80.07%**

Continued

Table 6.1: *Continued*

Variable	Sample Proportion (Mean)			Significance Test ^b			Row % Enroll=1 ^c
	Column % Full Sample N=1,041 ^a	Column % Enroll=0	Column % Enroll=1	χ^2 (t) Test Statistic	df	p- value	
<i>Sex</i>				15.95	1		
Male	48.98%	57.39%	44.42%	8.14	1	<.01	58.84%**
Female	51.02%	42.61%	55.58%	7.81	1	<.01	70.66%**
<i>Race/Ethnicity</i>				24.24	3	<.01	
White (Base)	73.86%	69.82%	76.05%	1.25	1	0.26	66.86%
Black	11.36%	13.48%	10.22%	2.22	1	0.14	58.34%
Hispanic	9.91%	14.55%	7.40%	12.24	1	<.01	48.43%**
Asian/Other	4.86%	2.15%	6.33%	8.53	1	<.01	85.04%**
ASVAB (percentile)	56.45%	39.94%	65.40%	-16.11 ^t	1039	<.01	
34.74 and below	24.98%	46.07%	13.56%	100.42	1	<.01	35.21%**
34.75 – 58.48	24.96%	29.13%	22.70%	3.93	1	0.05	59.00%*
58.48 – 80.05	25.00%	17.57%	29.03%	12.46	1	<.01	75.32%**
80.06 and over	25.07%	7.24%	34.72%	71.49	1	<.01	89.86%**
Ever Smoked				19.09	1	<.01	
No	47.56%	38.37%	52.54%	10.01	1	<.01	71.66%**
Yes	52.44%	61.63%	47.46%	9.08	1	<.01	58.72%**
Probability of Degree	79.00%	64.07%	87.09%	-14.32 ^t	1039	<.01	
0% – 49%	10.42%	23.35%	3.41%	90.61	1	<.01	21.24%**
50%	31.07%	39.75%	26.37%	13.65	1	<.01	55.07%**
51% – 99%	20.85%	13.84%	24.65%	13.30	1	<.01	76.69%**
100%	37.66%	23.06%	45.57%	31.90	1	<.01	78.49%**
CENSUS REGION CONTROLS				1.61	3	0.66	
Midwest	33.91%	32.90%	34.35%	-	-	-	61.50%
Northeast	14.90%	16.55%	14.01%	-	-	-	65.91%
South	31.04%	31.55%	30.76%	-	-	-	64.30%
West	20.15%	19.00%	20.78%	-	-	-	66.88%
RISK PREFERENCES							
Relative Risk Tolerance	0.9663	0.8227	1.0441	-3.99 ^t	1039	<.01	
0.331 and below	25.00%	27.88%	23.44%	1.87	1	0.17	60.83%**
0.332 – 0.698	25.07%	35.33%	19.52%	23.64	1	<.01	50.51%**
0.699 – 1.028	25.02%	18.19%	28.72%	10.52	1	<.01	74.47%**
1.029 and over	24.91%	18.60%	28.32%	8.99	1	<.01	73.77%**

Source: Weighted proportions and sample means of a restricted sample of the NLSY97.

^a The reported *N* reflects the actual (i.e., unweighted) number of cases in the full sample.

^b χ^2 tests of homogeneity were used to test for significance differences of proportions between categories. *t* tests were used for continuous variables and are indicated with a superscript *t*. See Appendix B for details.

^c Significance test is a binomial proportion test where the null hypothesis is 64.87%, the full sample proportion of those who enroll. See Appendix B for details.

[†] p<.10, *p<.05, **p<.01

Table 6.2. Relative Risk Tolerance at Interview Date

Risk Category	Risk Tolerance					<i>N</i> ^a
	Mean	SD	Minimum	Median	Maximum	
Women						
1	0.217	0.002	0.213	0.217	0.222	296
2	0.478	0.005	0.468	0.478	0.485	121
3	0.688	0.006	0.676	0.689	0.698	69
4	1.298	0.020	1.244	1.301	1.330	61
All	0.454	0.339	0.213	0.221	1.330	547
Men						
1	0.269	0.004	0.258	0.269	0.277	218
2	0.593	0.007	0.581	0.593	0.607	137
3	0.855	0.010	0.838	0.853	0.874	49
4	1.959	0.036	1.891	1.960	2.030	90
All	0.743	0.621	0.258	0.589	2.030	494

Source: Weighted, restricted sample of the NLSY97.

^a The reported *N* reflect the actual (i.e., unweighted) number of observations.

Table 6.3. Relative Risk Tolerance at Interview Date from Light and Ahn (2010)

Risk Category	Risk Tolerance					N
	Mean	SD	Minimum	Median	Maximum	
Women						
1	0.25	0.18	0.11	0.20	1.06	1,601
2	0.46	0.25	0.21	0.42	1.43	430
3	0.60	0.30	0.26	0.55	1.76	556
4	1.15	0.85	0.38	0.99	3.67	628
All	0.51	0.56	0.11	0.38	3.67	3,214
Men						
1	0.32	0.26	0.11	0.26	1.31	1,480
2	0.59	0.37	0.23	0.49	1.80	366
3	0.74	0.44	0.29	0.60	2.25	582
4	1.65	1.30	0.43	1.28	5.42	870
All	0.77	0.90	0.11	0.47	5.42	3,298

Source: NLSY79, Light and Ahn (2010), p. 905.

6.2 Multivariate Results

This section presents the findings of the logistic regressions described in Sections 4.2 and 5.4. Section 6.2.1 discusses the findings of Model 1 and Section 6.2.2 discusses the findings regarding differences between sexes.

6.2.1 Model 1

For the primary analysis in this study, Model 1, logistic regression was used to examine the effect of explanatory variables on the likelihood of higher education enrollment by age 20. Table 6.4 presents the results of Model 1, including parameter estimates and corresponding standard errors, *p*-values, and odds ratios. Model fit statistics can also be found in Table 6.4. Table 6.5 presents the computed average marginal effects from the estimates of Model 1. The following discussion refers primarily to the average marginal effects from Table 6.5 and considers results significant if $p < .05$ unless otherwise noted.

Generally, the model explained a substantial proportion of the variation in enrollment decisions. Model 1 had a Cox and Snell pseudo *r*-squared value of .347 and a Nagelkerke pseudo *r*-squared value of .473. Model 1 had a *c* statistic of 0.860, which measures the area under the Receiver Operating Characteristic (ROC) curve (Allison, 2012, pp. 68-80). ROC curves graph the sensitivity (the proportion of *Enroll*=1 that were correctly predicted) and the specificity (the proportion of *Enroll*=0 that were correctly predicted). The greater the *c* statistic the better the predictive power of the model (Allison, 2012).

High school GPA was significantly and positively associated with higher education enrollment. Increasing GPA by one unit led to a nearly 10 percentage point increase in the likelihood of enrolling in higher education. Young adults who lived in other living situations did not differ significantly from those who lived with both biological parents with one exception; young adults who lived with biological fathers only were nearly 19 percentage points less likely to enroll in higher education than respondents who lived with both biological parents.

Parental education was an important predictor of enrollment decisions. Compared to respondents whose parents' highest education level was high school, those who attended less than four years, four years, and more than four years of college were 12, 11, and 15 percentage points more likely, respectively, to enroll in higher education. Respondents whose parents had dropped out of high school were not significantly different than those who finished high school.

Family net worth was a somewhat significant, but not meaningful, predictor of enrollment ($p < .10$). An increase in net worth of \$100,000 was associated with a two percentage point increase in the likelihood of enrollment. Family income was positively related to enrollment, although the effect was also small – increasing income by \$10,000 would result in a one percentage point increase in the likelihood of enrollment. Parental estimates of the probability that the respondent would have a college degree were significantly related to enrollment but, again, the effect was small – increasing the probability by 10 percentage points was associated with an increase in the likelihood of enrollment by about two percentage points.

Gender was also a significant predictor of enrollment; women were about 9 percentage points more likely to enroll in higher education than men. Compared to White young adults, respondents of other races/ethnicities were generally not more or less likely to enroll. The one exception was that Black respondents were about 7 percentage points more likely to enroll than White respondents ($p < .10$). Cognitive ability was significantly and positively related to enrollment – increasing the ASVAB score by 10 percentage points was associated with a three percentage point increase in the likelihood of enrollment. Young adults who had ever smoked were about 6 percentage points less likely to enroll in higher education compared to young adults who had never smoked. The subjective probability of college completion was a significant predictor of college enrollment; however, the effect of was somewhat small. Increasing the probability of completion by 10 percent was associated with a two percentage point increase in the probability of enrollment.

Generally, as in the descriptive analysis, there were not significant differences in the probability of enrollment between the four census regions. However, young adults living in the West census region were about 7 percentage points more likely to enroll in higher education compared to those living in the Midwest ($p < .10$).

To interpret the effect of relative risk tolerance on the probability of enrollment, it is most helpful to compute the mean predicted probabilities at different levels of the age-adjusted relative risk tolerance. Generally speaking, there was a clear, positive relationship between risk tolerance and the mean predicted probability of enrollment. Table 6.6 presents the mean predicted probabilities separately for women and men since gender was an important predictor of enrollment and women have lower levels of risk

tolerance, on average, compared to men. For women, young adults in the first (bottom) quartile had a mean predicted probability of 57 percent while those in the fourth (top) quartile had a mean predicted probability of 78 percent. For men, young adults in the first (bottom) quartile had a mean predicted probability of 42 percent while those in the fourth (top) quartile had a mean predicted probability of 73 percent. Table 6.7 presents the mean predicted probability of enrollment by the risk tolerance categories created from the income gamble question and shows a similar pattern. The predicted probability for those in the lowest risk tolerance category was 54 percent while those in the highest risk tolerance category had a predicted probability of 77 percent.

Table 6.4. Model 1: Logistic Regression of Higher Education Enrollment by Age 20.

Variable^a <i>N</i> =1,041 ^b	Estimate (SE)	Wald χ^2	<i>p</i>- value	Odds Ratio	95% Wald C.I. Odds Ratio	
Intercept	-5.831 (0.561)	108.127	<.001	-	-	-
INDIVIDUAL-LEVEL CONTROLS						
High School GPA	0.675 (0.149)	20.636	<.001	1.963	1.467	2.627
<i>Family Structure</i> (Both Biological Parents)						
Two Parents, One Biological	-0.157 (0.250)	0.393	0.531	0.855	0.524	1.395
Biological Mother Only	-0.000 (0.214)	0.000	0.999	1.000	0.658	1.520
Biological Father Only	-1.271 (0.512)	6.162	0.013	0.281	0.103	0.765
Other Parental Figures	-0.303 (0.474)	0.411	0.522	0.738	0.292	1.868
<i>Parental Education</i> (12 th Grade)						
Less Than 12th Grade	0.182 (0.268)	0.462	0.497	1.199	0.710	2.027
Some College	0.815 (0.209)	15.165	<.001	2.260	1.499	3.407
Four Years of College	0.751 (0.280)	7.201	0.007	2.119	1.224	3.668
More than Four Years	1.032 (0.343)	9.067	0.003	2.806	1.433	5.491
Household Net Worth (\$10,000 increments)	0.015 (0.008)	3.259	0.071	1.015	0.999	1.032
Household Income (\$10,000 increments)	0.076 (0.036)	4.451	0.035	1.079	1.005	1.157
Probability of Degree – Parent ^c	0.107 (0.033)	10.718	0.001	1.113	1.044	1.186
<i>Sex</i> (Male)						
Female	0.598 (0.177)	11.406	0.001	1.818	1.285	2.572
<i>Race/Ethnicity</i> (White)						
Black	0.482 (0.246)	3.831	0.050	1.619	0.999	2.623
Hispanic	-0.091 (0.252)	0.131	0.718	0.913	0.557	1.497

Continued

Table 6.4: *Continued*

Variable^a <i>N</i> =1,041 ^b	Estimate (SE)	Wald χ^2	<i>p</i>- value	Odds Ratio	95% Wald C.I. Odds Ratio	
Asian/Other	0.724 (0.534)	1.838	0.175	2.062	0.724	5.868
ASVAB ^c	0.216 (0.038)	31.529	<.001	1.241	1.151	1.338
Ever Smoked	-0.383 (0.173)	4.886	0.027	0.682	0.485	0.957
Probability of Degree ^c	0.136 (0.033)	16.467	<.001	1.145	1.073	1.223
CENSUS REGION CONTROLS (Midwest)						
Northeast	0.066 (0.270)	0.060	0.806	1.069	0.629	1.816
South	0.027 (0.218)	0.016	0.901	1.028	0.671	1.574
West	0.449 (0.256)	3.069	0.080	1.566	0.948	2.587
RISK PREFERENCES						
Relative Risk Tolerance	0.412 (0.107)	14.984	<.001	1.511	1.226	1.861
MODEL STATISTICS						
-2Log-Likelihood	935.748		<i>c</i> Statistic		0.860	
AIC	982.803		Cox & Snell Pseudo R ²		.347	
Percent Concordant	85.90		Nagelkerke Pseudo R ²		.473	

Source: Unweighted logistic regression of a restricted sample of complete cases in the NLSY97.

^a Reference categories are in parentheses.

^b The reported *N* reflect the actual (i.e., unweighted) number of observations.

^c Increments of 10 percentage points.

Table 6.5. Model 1: Average Marginal Effects

Variable ^a N=1,041 ^b	Estimate	SE	z	p-value	95% CI	
INDIVIDUAL-LEVEL CONTROLS						
High School GPA	0.099	0.021	4.710	0.000	0.058	0.140
<i>Family Structure (Both Biological Parents)</i>						
Two Parents, One Biological	-0.023	0.037	-0.630	0.530	-0.095	0.049
Biological Mother Only	0.000	0.031	0.000	0.999	-0.061	0.061
Biological Father Only	-0.186	0.074	-2.510	0.012	-0.332	-0.041
Other Parental Figures	-0.044	0.069	-0.640	0.521	-0.180	0.091
<i>Parental Education (12th Grade)</i>						
Less Than 12th Grade	0.027	0.039	0.680	0.496	-0.050	0.103
Some College	0.119	0.030	4.000	0.000	0.061	0.178
Four Years of College	0.110	0.040	2.720	0.007	0.031	0.189
More than Four Years	0.151	0.049	3.060	0.002	0.054	0.248
Household Net Worth						
(\$10,000 increments)	0.002	0.001	1.820	0.070	0.000	0.005
Household Income						
(\$10,000 increments)	0.011	0.005	2.120	0.034	0.001	0.021
Probability of Degree – Parent ^c	0.016	0.005	3.340	0.001	0.006	0.025
<i>Sex (Male)</i>						
Female	0.088	0.025	3.440	0.001	0.038	0.137
<i>Race/Ethnicity (White)</i>						
Black	0.071	0.036	1.970	0.049	0.000	0.141
Hispanic	-0.013	0.037	-0.360	0.718	-0.086	0.059
Asian/Other	0.106	0.078	1.360	0.174	-0.047	0.259
ASVAB ^c	0.032	0.005	5.930	0.000	0.021	0.042
Ever Smoked (No)						
Yes	-0.056	0.025	-2.230	0.026	-0.106	-0.007
Probability of Degree ^c	0.020	0.005	4.180	0.000	0.011	0.029
CENSUS REGION CONTROLS (Midwest)						
Northeast	0.010	0.040	0.250	0.806	-0.068	0.087
South	0.004	0.032	0.120	0.901	-0.058	0.066
West	0.066	0.037	1.760	0.078	-0.007	0.139
RISK PREFERENCES						
Relative Risk Tolerance	0.060	0.015	3.970	0.000	0.031	0.090

Source: Unweighted logistic regression of a restricted sample of complete cases in the NLSY97.

^a Reference categories are in parentheses.

^b The reported *N* reflect the actual (i.e., unweighted) number of observations.

^c Increments of 10 percentage points.

Table 6.6. Mean Predicted Probability of Enrollment by Quartile of RRT

Quartile of Relative Risk Tolerance	Mean Predicted Probability of Enrollment	<i>N</i>
Women		547
First	.5744	137
Second	.6409	140
Third	.7174	129
Fourth	.7804	141
Men		494
First	.4199	135
Second	.5071	122
Third	.6256	119
Fourth	.7288	118

Source: Unweighted logistic regression of a restricted sample of complete cases in the NLSY97.

Table 6.7. Mean Predicted Probability of Enrollment by RRT Category

Risk Tolerance Categories	Mean Predicted Probability of Enrollment	<i>N</i>
1	.5454	514
2	.6558	258
3	.7155	118
4	.7683	151

Source: Unweighted logistic regression of a restricted sample of complete cases in the NLSY97.

6.2.2 Gender Differences

Given the overwhelming evidence from the literature regarding differences between men and women on higher education choices, labor market outcomes, and risk preferences, a likelihood ratio test was used to examine whether or not men and women needed to be analyzed separately to answer the research questions. Gender may plausibly affect each parameter in the model, suggesting the need for a fully interacted model. The results of the likelihood ratio test are shown in Table 6.8. The conclusion of the test is

that the restricted model, Model 1, is a better choice than the two separate models. A brief summary of the separate models is shown in Table 6.9 to illustrate the impact of running separate equations for men and women on the key parameter of interest. Specifically, there do appear to be some differences between the two sexes. For example, the parameter estimate for relative risk tolerance for women is nearly twice the size of the estimate in Model 1 and nearly three times as large as the parameter estimate for males. However, the conclusion regarding the research question is not different than the conclusion reached using Model 1.

Table 6.8. Likelihood Ratio Test: Pooled versus Separate Logits for Men and Women

Model	-2LogLikelihood	df	p-value
Restricted (pooled) Model – Model 1	934.803	23	-
Unrestricted Model	913.166	44	-
Model 1.M	462.372	22	-
Model 1.W	450.794	22	-
Likelihood Ratio Test Statistic, χ^2	21.637	21	0.421
Conclusion	Fail to reject null hypothesis: Restricted Model is more appropriate		

Table 6.9. Gender Differences and Key Parameters

Variable^a	Model 1 <i>N</i> =1,041 ^b	Model 1.M <i>N</i> =494 ^b	Model 1.W <i>N</i> =547 ^b
Parameter Estimates			
	(SE)		
Relative Risk Tolerance	0.412*** (0.107)	0.331** (0.117)	0.815** (0.264)
Model Statistics			
-2LogLikelihood	934.803	462.372	450.794
AIC	982.803	508.372	497.974
Percent Concordant	85.90	85.10	87.30
<i>c</i> Statistic	0.860	0.851	0.874
Cox & Snell Pseudo R ²	0.347	0.352	0.351
Nagelkerke Pseudo R ²	0.473	0.472	0.491

Source: Unweighted logistic regressions of a restricted sample of complete cases in the NLSY97.

^a Full regression output is available from the author.

^b The reported *N* reflect the actual (i.e., unweighted) number of observations.

[†] *p*<.10, **p*<.05, ***p*<.01, *** *p*<.001

6.3 Sensitivity Analysis Results

As discussed in Section 5.5, sensitivity analysis was performed to examine the robustness of the findings in Model 1 to empirical specifications. This section presents the results of these robustness checks with respect to missing data (Section 6.3.1), measurement of risk tolerance (Section 6.3.2), and timing of variable measurement (Section 6.3.3).

6.3.1 Missing Data

The first missing data analysis involved separating the sample into complete cases and incomplete cases and examining the descriptive frequencies of the explanatory variables. Table 6.10 presents the number missing of each variable and the descriptive

differences between the full potential sample, the cases that were excluded from the primary analysis (selected=0), and the cases that were used for the primary analysis (selected=1).

In terms of missing data, many of the variables were not missing observations or were missing a small number. However, there were six variables that were missing at least 400 observations: (1) high school GPA, (2) family net worth, (3) family income, (4) parent estimate of college completion, (5) cognitive ability, and (6) risk tolerance. Given the large number of missing observations of these variables, sample proportions and means are not reported in Table 6.10. Since parental education, race/ethnicity, and census region were missing few or no observations, it is useful to examine differences in these proportions between the full sample and the sample that was selected.

The complete cases had a slightly smaller proportion of parents with a high school diploma or less compared to the full sample. Forty-two percent of the full sample had parents who had a high school diploma or less compared to 40 percent of the selected sample. The selected sample also had a higher proportion of Whites (74 percent) and lower proportions of Blacks (11 percent) and Hispanics (10 percent) than the full sample (69 percent, 13 percent, and 12 percent, respectively). The selected sample also had a lower proportion of respondents in the Northeast region (15 percent) compared to the full sample (19 percent). In sum, the cases that did not have missing observations tended to be more educated, to be White, and to be located in the Midwest rather than the Northeast compared to the full sample. However, these differences tended to be relatively small (less than 5 percent).

Table 6.10. Missing Data Analysis: Descriptive Statistics By Selected

Variable	Number Missing ^a	Sample Proportion/Mean		
		Full Sample N=2,781 ^a	Incomplete (Selected=0)	Complete (Selected=1)
DV: Enroll	0	62.03%	60.21%	64.87%
INDIVIDUAL-LEVEL CONTROLS				
High School GPA	489	— ^b	— ^b	— ^b
<i>Family Structure</i>	5			
Both Biological Parents	-	56.07%	55.78%	56.53%
Two Parents, One Biological	-	13.94%	13.81%	14.13%
Biological Mother Only	-	22.27%	21.59%	23.32%
Biological Father Only	-	3.33%	3.39%	3.23%
Other Parental Figures	-	4.40%	5.43%	2.79%
<i>Parental Education</i>	0			
Less Than 12th Grade	-	14.19%	16.53%	10.54%
12 th Grade	-	28.04%	27.19%	29.37%
Some College	-	26.41%	24.87%	28.83%
Four Years of College	-	16.85%	16.74%	17.03%
More than Four Years	-	14.50%	14.66%	14.24%
Household Net Worth (\$)	764	— ^b	— ^b	— ^b
Household Income	762	— ^b	— ^b	— ^b
Probability of Degree- Parent	422	— ^b	— ^b	— ^b
<i>Sex</i>	0			
Male	-	50.08%	50.79%	48.98%
Female	-	49.92%	49.21%	51.02%
<i>Race/Ethnicity</i>	0			
White	-	69.37%	66.42%	74.00%
Black	-	13.26%	14.47%	11.36%
Hispanic	-	11.92%	13.20%	9.91%
Asian/Other	-	4.67%	5.02%	4.11%
ASVAB	507	— ^b	— ^b	— ^b
Ever Smoked	6	51.85%	51.31%	52.44%
Probability of Degree	29	77.90%	77.18%	79.00%
CENSUS REGION CONTROLS				
Midwest	-	27.99%	24.22%	33.91%
Northeast	-	18.92%	21.48%	14.90%
South	-	32.82%	33.96%	31.04%
West	-	20.27%	20.34%	20.15%
RISK PREFERENCES				
Relative Risk Tolerance	462	— ^b	— ^b	— ^b

Source: Weighted proportions and sample means of a restricted sample of the NLSY97. Selected cases were complete cases that were used in Model 1; cases with any missing observations (Selected=0) were not included in Model 1.

^a The reported Number Missing and *N* reflect the actual (i.e., unweighted) number of observations.

^b Proportions and means were omitted for variables missing a large number of observations.

The sensitivity of the findings to the inclusion of the missing cases in the logistic regression was also tested using mean imputation and dummy variables to indicate whether or not a variable was missing. This procedure nearly doubled the sample size to 1,969. Table 6.11 reports the estimates for relative risk tolerance and presents the results of the imputed logistic regression, Model 1.IMP. Table 6.11 indicates that the dummy variables for missing values of GPA, net worth, and income were not statistically significant; however, the dummy variable for ASVAB missing was significant. This indicates that those who chose not to take the ASVAB were systematically different than those who did take the exam. Specifically, those who opted out of the ASVAB were significantly less likely to enroll in higher education than those who did take the exam. The estimated coefficient of relative risk tolerance was affected by the additional cases, although the effect was still significant.

Table 6.11. Sensitivity to Missing Data Using Imputation

	Model 1 <i>N</i> =1,041 ^b	Model 1.IMP <i>N</i> =1,969 ^b
Missing Data Imputed	No	Yes ^c
Variable^a	Parameter Estimates (SE)	
GPA: Missing	-	-0.155 (0.151)
NW: Missing	-	0.024 (0.166)
INC: Missing	-	0.169 (0.163)
ASVAB: Missing	-	-10.083 ^{***} (1.328)
Relative Risk Tolerance	0.413 ^{***} (0.107)	0.192 ^{**} (0.073)
Model Statistics		
-2LogLikelihood	935.748	1869.358
AIC	983.748	1925.358
Percent Concordant	85.90	84.30
<i>c</i> Statistic	0.860	0.844
Cox & Snell Pseudo R ²	0.346	0.321
Nagelkerke Pseudo R ²	0.472	0.436

Source: Unweighted logistic regressions of a restricted sample of complete cases in the NLSY97.

^a Full regression output is available from the author.

^b The reported *N* reflect the actual (i.e., unweighted) number of observations.

^c Mean imputation used for *hsGPA*, *FamNW*, *FamInc*, *ASVAB*. Corresponding dummy variables for each of the four imputed variables were used to distinguish imputed cases (code 1) from observed cases (coded 0).

[†] $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$

6.3.2 Measurement of Risk Tolerance

The sensitivity of the findings to the measurement of risk tolerance was also tested by using alternate measures of time-invariant risk preferences. Model 1 was run using four alternate measures of risk tolerance: risk tolerance categories from the income gamble questions (Model 2), the general willingness to take risks scale (Model 3), willingness to take risks in financial matters scale (Model 4), and willingness to take risks

in major life changes scale (Model 5). Table 6.12 presents the results of each model with respect to the key parameters of interest. Response rates were slightly higher on the willingness to take risks scale questions so the number in each model varies slightly. Table 6.12 shows that the effect of risk preferences is significant and positive across all models.

Table 6.12. Sensitivity to Measurement of Risk Tolerance

Variable^a	Model 1 <i>N</i> =1,041 ^b	Model 2 <i>N</i> =1,041 ^b	Model 3 <i>N</i> =1,050 ^b	Model 4 <i>N</i> =1,050 ^b	Model 5 <i>N</i> =1,044 ^b
Risk Tolerance Measure	Relative Risk Tolerance	Risk Tolerance Categories	<i>GenRisk</i>	<i>FinRisk</i>	<i>ChgRisk</i>
Parameter Estimates (SE)					
Risk Tolerance Measure (varies by model)	0.412 ^{***} (0.107)	- ^c	0.128 ^{***} (0.034)	0.078 [*] (0.032)	0.075 [*] (0.032)
RRT2	-	0.507 [*] (0.207)	-	-	-
RRT3	-	0.396 (0.278)	-	-	-
RRT4	-	1.086 ^{***} (0.268)	-	-	-
Model Statistics					
-2LogLikelihood	935.748	930.762	932.273	944.103	938.928
AIC	983.748	982.762	980.273	992.103	986.928
Percent Concordant	85.90	86.00	86.40	85.90	85.90
<i>c</i> Statistic	0.860	0.861	0.864	.0860	0.859
Cox & Snell Pseudo R ²	0.346	0.349	0.354	0.347	0.346
Nagelkerke Pseudo R ²	0.472	0.476	0.483	0.473	0.471

Source: Unweighted logistic regressions of a restricted sample of complete cases in the NLSY97.

^a Full regression output is available from the author.

^b The reported *N* reflect the actual (i.e., unweighted) number of observations.

^c Omitted category is RRT1.

[†] p<.10, *p<.05, **p<.01, *** p<.001

6.3.3 Timing of Variable Measurement

Lastly, the sensitivity of the findings to the timing of variable measurement was tested by running Model 1 separately for 15 year olds and 16 year olds and using a likelihood ratio test to see if the coefficients were significantly different by age. This test allows for each parameter estimate to vary by age of respondent. Table 6.13 shows that the null hypothesis was rejected, suggesting that the restricted model is a better fit than the two separate models.

In order to determine how sensitive the results were to the timing of variable measurement, Table 6.14 presents the key parameter estimates of the separate logits versus the pooled logit (Model 1). Approximately 54 percent of respondents were younger than age 16 and the remaining 46 percent were 16 or older. In terms of the model fit statistics, the model does a relatively better job explaining the variation in higher education choices of the older sample than the younger sample. However, the effect of relative risk tolerance is positive and significant ($p < .10$ in the case of relative risk tolerances in Model 1.O).

Table 6.13. Likelihood Ratio Test: Pooled versus Separate Logits for Respondents Younger than Age 16 and Respondents Age 16 or Older.

Model	-2LogLikelihood	df	p-value
Restricted (pooled) Model – Model 1	934.803	23	-
Unrestricted Model	916.085	46	-
Younger (Age<16) – Model 1.Y	505.671	23	-
Older (Age≥16) – Model 1.O	410.324	23	-
Likelihood Ratio Test Statistic, χ^2	18.718	23	0.712
Conclusion	Fail to reject null hypothesis: Restricted Model is more appropriate		

Table 6.14. Sensitivity to Timing of Variable Measurement

Variable ^a	Model 1 N=1,041 ^b	Model 1.Y (Age<16) N=561 ^b	Model 1.O (Age≥16) N=480 ^b
Parameter Estimates (SE)			
Relative Risk Tolerance	0.412*** (0.107)	0.256 [†] (0.136)	0.681*** (0.186)
Model Statistics			
-2LogLikelihood	934.803	505.671	410.324
AIC	982.803	553.671	458.324
Percent Concordant	85.90	85.50	87.60
<i>c</i> Statistic	0.860	0.855	0.877
Cox & Snell Pseudo R ²	0.347	0.332	0.387
Nagelkerke Pseudo R ²	0.473	0.456	0.522

Source: Unweighted logistic regressions of a restricted sample of complete cases in the NLSY97.

^a Full regression output is available from the author.

^b The reported *N* reflect the actual (i.e., unweighted) number of observations.

[†] p<.10, *p<.05, **p<.01, *** p<.001

Chapter 7: Discussion

This chapter provides a discussion regarding the findings presented in Chapter 6. Section 7.1 discusses the current study's findings, compares the results to previous research, and details the contribution to the literature on consumer higher education enrollment decisions. Section 7.2 addresses the limitations regarding the research methodology. Section 7.3 discusses the implications of the results and provides questions and directions for future research. Lastly, Section 7.4 concludes the discussion with a brief summary of the primary contributions of the current study.

7.1 Interpretation and Evaluation of Results

As discussed in Section 1.2, the current study sought to determine the effect of consumer risk preferences on higher education enrollment decisions. Based on the theoretical model in Section 4.1, consumer risk preferences were hypothesized to have an effect on higher education enrollment – the findings presented in Chapter 6 provide evidence in support of the research hypothesis. Even after controlling for a number of individual characteristics, including time preferences and subjective risk perceptions, consumer risk preferences had a robust and significant effect on enrollment. Specifically, higher risk tolerance was associated with greater likelihoods of enrollment.

This analysis is unique in that it was the first to examine the effect of risk preferences on higher education enrollment choices while accounting for subjective probabilities of college completion and time preferences. This study supports the small but growing literature (Belzil & Leonardi, 2009; Brown et al., 2006; Hartog & Diaz-Serrano, 2007) that shows consumer risk preferences are significant predictors of higher education choices. The sensitivity analysis demonstrates that risk tolerance had a robust and positive effect on the likelihood of enrollment.

While much of the popular media coverage surrounding student loans has suggested that consumers are making poor (i.e., irrational) decisions, the risky nature of the decision has been largely ignored. At a fundamental level, the higher education choices of young adults may be inferred to be rational since consumers perceive that human capital investment is a risky choice and individual risk preferences do in fact have an impact on enrollment decisions. However, beyond this very basic level, further analyses are needed to fully appreciate the rationality, or lack thereof, of young adults' decision making process in higher education choices.

The finding that more risk tolerant individuals are more likely to enroll in higher education implies that, to consumers, attending higher education may seem more risky than entering the labor market directly. Although researchers have documented the lower unemployment rates that college graduates face compared to those with a high school diploma (Oreopoulos & Salvanes, 2011), from the consumer's perspective, there are a greater number of uncertainties and decision points in pursuing higher education. Choosing to enter the labor market directly is a comparatively simpler choice with less uncertainty. Although there may be greater risk of unemployment without a college

degree, consumers may choose to not enroll due to the perceived complexity (e.g., where to go, what major to pursue, how to finance, etc.) and the larger number of uncertain outcomes (e.g., ability to graduate, economic conditions at graduation, etc.).

This study also has found supporting evidence of previous findings in the college going literature. For example, the current analysis supports previous findings that student ability and family background (e.g., Belzil & Hansen, 2003) are important predictors of college enrollment. Academic ability, as measured by high school GPA, and cognitive ability, as measured by the ASVAB percentile score, were both highly significant positive predictors of enrollment. Parental education was also an important predictor of enrollment. Young adults living with parents who had attended at least one year of college were between 11 and 15 percentage points more likely to enroll in higher education.

While Lillard and Gerner (1999) found that students from disrupted families were less likely to enroll in higher education, the current study found that young adults from disrupted families were generally not significantly different than young adults who lived with both biological parents. The noteworthy exception to this generalization is that young adults who lived with the biological father only were significantly (19 percentage points) less likely to enroll in higher education.

The findings also support the literature (e.g., DiPrete & Buchmann, 2013) that women are more likely to enroll in higher education than men. This may be explained by the earnings gap in the labor market – women may enroll at higher rates in an effort to close this gap. The same logic may also apply to the findings regarding racial/ethnic groups. Although the descriptive analysis showed that Blacks have lower enrollment rates

than Whites, the multivariate analysis showed that Blacks were actually 7 percentage points more likely to enroll than Whites. Black young adults may be aware of the earnings gap in the labor market, so assuming all other covariates are held constant, they may be more willing than Whites to enroll in higher education in an effort to reduce this gap.

7.2 Limitations

This section addresses several limitations regarding the current study. As previously discussed in Section 5.3, one challenge in studying the higher education decision making process of young adults is that there is not an exact decision point in which enrollment decisions are made once and for all. Simplifying assumptions had to be made in order to carry out the analysis, but as with any model, this is hardly the reality that consumers face.

As with any large, secondary dataset, there were also limitations in terms of measuring key constructs. The most important limitation to point out was the lack of a direct measure of time preferences. As previous research has suggested (Hartog & Diaz-Serrano, 2007), empirical research in this area needs to find ways to disentangle the effects of time preferences and risk preferences. To this end, the current study had to rely on a proxy, smoking behavior, as a measure of time preference. Although definitive statements about the effect of time preference cannot be made, the level of individual controls in the model suggests that other plausible explanations for the effect of smoking may be ruled out (e.g., as a measure of socio-economic status).

Although the NLSY97 contains a rich variety of data on young adults, this data is collected at a variety of points in time. For example, the parent survey primarily took place in the first survey round, but ideally, data would have been collected from the parents as long as students lived at home. The expectation question regarding the likelihood of college completion was also not asked of the entire sample, which limits the sample size that can be used for research questions similar to the current study. Since key variables may only be asked at specific time points, the timing issue is an important one that researchers must consider in determining appropriate methodologies.

In terms of the generalizability of the findings, the sensitivity analysis revealed that the sample of complete cases analyzed in the current study may be slightly skewed towards more educated, White households living in the Midwest compared to the general population. Although the robustness checks indicated that the primary research finding is unlikely to be altered by this limitation, the parameter estimates should be viewed somewhat cautiously since the sample may not be representative.

Despite these limitations, the primary conclusion of the current study, that risk tolerance is significantly and positively related to the likelihood of enrollment, is unlikely to be negated. The sensitivity analysis provides evidence that the finding is robust to the following empirical specifications: sample selection (i.e., complete cases only), measurement of risk tolerance, and the timing of variable measurement.

7.3 Implications and Future Research

The primary implication of this study is that more research is needed to fully explore consumer decisions in human capital investment. Specifically, researchers need

to investigate how consumers perceive the risk involved in higher education investment and the extent to which incorrect risk perceptions are negatively impacting consumer well-being. Young adults may be discouraged from attending higher education based on inflated risk perceptions and misinformed views of the costs and benefits. Evidence from behavioral economics and judgment and decision making suggests that it is quite likely that young adult perceptions regarding the riskiness of college choices are not accurate. For example, literature shows that students are prone to overestimate both the costs and benefits of higher education (Avery & Kane, 2004). These inaccuracies may lead to sub-optimal human capital investment decisions.

Since there is not a single decision point in which consumers decide to pursue higher education, future research should utilize qualitative and mixed method approaches to further examine this decision making process as a risky decision. Interviews and experimental methods would shed more light on how the riskiness of human capital investment is perceived and how higher education decisions are made. There may be productive ways to improve consumer decision making through the use of decisions aids that help consumers make informed choices. The recent literature examining the option value of higher education (e.g., Jacobs, 2007) is also a productive area for further research since consumers may delay entry into higher education in order to reduce uncertainty.

Further research should investigate the degree to which college expectations match college attainment by investigating the accuracy of student estimates of college completion. This may have very interesting implications for policies designed to encourage college attendance. Furthermore, the sensitivity analysis revealed that the

general willingness to take risks scale did a relatively better job of predicting enrollment than the income gamble question. This may imply that consumers may find the simple willingness to take risks scales easier to understand than the income gamble questions. Future research may further investigate these various risk measures and their predictive power of risky behavior.

The risky nature of higher education choices is largely ignored in popular media accounts and policy conversations. For example, although the recent activity surrounding gainful employment legislation (see Fain, 2014) has focused on for-profit institutions, many fear that these rules may eventually be applied to traditional education. The current study suggests that policies of this nature need to consider the uncertainty and risk involved in pursuing higher education. Many outcomes, especially those in the labor market, are beyond the institution's and student's control. There may be more effective policy tools that can be used to hold institutions accountable while also recognizing the risk involved.

For researchers interested in using the NLSY97 to examine higher education choices, this study has highlighted a potentially serious sample selection issue if not considered. Sample selection bias may be introduced to multivariate analyses that intend to use the ASVAB score as a proxy for cognitive ability if respondents who choose to not take the exam are excluded from analyses. These respondents may simply be averse to taking exams and/or they may have an average cognitive ability lower than the general population. Future research should investigate this issue more thoroughly.

Gender differences are also worth exploring in greater detail. Although the fully interacted model was rejected, there were differences between the two logits run

separately for men and women, suggesting that more carefully selected interaction terms may be important additions to the model. Furthermore, since Perna (2006) suggests that habitus is an important construct in the decision to attend higher education, future research should test for group differences by race and SES. While enrollment was an appropriate dependent variable to analyze the decision making process, as discussed by Turner (2004), future research might consider using attainment measures (e.g., years of education or credit hours) as the dependent variable for studies regarding human capital accumulation.

The current study focused on traditional college students, but many in higher education have acknowledged that traditional students make up a decreasing share of students on college campuses (see Laitinen, 2012; Selingo, 2014). Therefore, future research should examine the decision making of non-traditional students. Furthermore, this study defined enrollment to include either two-year or four-year schools. There may be interesting differences between these students and risk perceptions and preferences that would be worth investigating. Many students also enroll concurrently between two-year and four-year schools (see Wang & Wickersham, 2013). Concurrent enrollment may be a risk management tool that consumers are using to test higher education or more accurately understand their own ability to succeed. Future research should explore these areas.

7.4 Conclusion

In conclusion, this research has shown that even after accounting for time preferences and subjective probabilities of college completion, risk preferences have a

significant and robust effect on higher education enrollment decisions. Specifically, the findings show that more risk tolerant young adults are significantly more likely to enroll in higher education. This finding has important implications for higher education stakeholders and identifies the risky nature of higher education choices as a key, but overlooked, area in the current conversation regarding student loans and the value of a college degree. A number of productive areas for future research have been identified. Ultimately, since human capital investment is a risky choice, decision making aids and ensuring informed choices may be promising avenues for providing assistance to consumers and for encouraging investment in human capital.

References

- Ahn, T. (2010). Attitudes toward risk and self-employment of young workers. *Labour Economics*, 17(2), 434-442. doi: 10.1016/j.labeco.2009.06.005
- Ajzen, I., & Fishbein, M. (1977). Attitude-behavior relations: A theoretical analysis and review of empirical research. *Psychological Bulletin*, 84(5), 888. doi: 10.1037/0033-2909.84.5.888
- Ajzen, I., Fishbein, M., Heilbroner, R. L., & Thurow, L. C. (1980). *Understanding attitudes and predicting social behavior*. Englewood Cliffs, NJ: Prentice-Hall.
- Allison, P. D. (2012). *Logistic regression using sas: Theory and application* (2nd ed.). Cary, NC: SAS Institute Inc.
- Andini, C. (2009). On the return risk link in education. *Applied Economics Letters*, 16(3), 307-314. doi: 10.1080/13504850601018395
- Andruska, E. A., Hogarth, J. M., Fletcher, C. N., Forbes, G. R., & Wohlgemuth, D. R. (2012). *Do you know what you owe? Students' understanding of their student loans*. Paper presented at the Academy of Financial Services, San Antonio, TX.
- Arcidiacono, P., Bayer, P., & Hizmo, A. (2010). Beyond signaling and human capital: Education and the revelation of ability. *American Economic Journal*, 2(4), 76-104. doi: 10.1257/app.2.4.76
- Ariely, D. (2008). *Predictably irrational: The hidden forces that shape our decisions*. New York, NY: HarperCollins.
- Arrow, K. J. (1971). *Essays in the theory of risk-bearing* (Vol. 2). Amsterdam: North-Holland Publishing Company.
- Athanassiadis, A. (2011). Economic returns and risks to investment in education: An application of the multifactor capm. *International Journal of Economic Sciences and Applied Research*(1), 95-120.
- Avery, C., & Hoxby, C. M. (2004). Do and should financial aid packages affect students' college choices? In C. M. Hoxby (Ed.), *College choices: The economics of where to go, when to go, and how to pay for it* (pp. 239-302). Chicago, IL: University of Chicago Press.

- Avery, C., & Kane, T. J. (2004). Student perceptions of college opportunities. The boston coach program. In C. M. Hoxby (Ed.), *College choices: The economics of where to go, when to go, and how to pay for it* (pp. 355-394). Chicago, IL: University of Chicago Press.
- Avery, C., & Turner, S. (2012). Student loans: Do college students borrow too much-or not enough? *Journal of Economic Perspectives*, 26(1), 165-192. doi: 10.1257/jep.26.1.165
- Barsky, R. B., Juster, F. T., Kimball, M. S., & Shapiro, M. D. (1997). Preference parameters and behavioral heterogeneity: An experimental approach in the health and retirement study. *The Quarterly Journal of Economics*, 112(2), 537-579. doi: 10.1162/003355397555280
- Baum, S., & O'Malley, M. (2003). College on credit: How borrowers perceive their education debt. *Journal of Student Financial Aid*, 33(3), 7-19.
- Bean, J. P. (1982). Conceptual models of student attrition: How theory can help the institutional researcher. *New Directions for Institutional Research*, 1982(36), 17-33. doi: 10.1002/ir.37019823604
- Bean, J. P. (1983). The application of a model of turnover in work organizations to the student attrition process. *Review of Higher Education*, 6(2), 129-148.
- Bean, J. P., & Metzner, B. S. (1985). A conceptual model of nontraditional undergraduate student attrition. *Review of Educational Research*, 55(4), 485-540. doi: 10.3102/00346543055004485
- Becker, G. S. (1962). Investment in human capital: A theoretical analysis. *Journal of Political Economy*, 70(5), 9-49. doi: 10.2307/1829103
- Becker, G. S. (1964). *Human capital: A theoretical and empirical analysis, with special reference to education*. Chicago, IL: University of Chicago Press.
- Belley, P., & Lochner, L. (2007). The changing role of family income and ability in determining educational achievement. *Journal of Human Capital*, 1(1), 37-89. doi: 10.1086/524674
- Belzil, C., & Hansen, J. (2003). Structural estimates of the intergenerational education correlation. *Journal of Applied Econometrics*, 18(6), 679-696. doi: 10.1002/jae.716
- Belzil, C., & Leonardi, M. (2007). Can risk aversion explain schooling attainments? Evidence from Italy. *Labour Economics*, 14(6), 957-970. doi: 10.1016/j.labeco.2007.06.005
- Belzil, C., & Leonardi, M. (2009). Risk aversion and schooling decisions.

- Bilkic, N., Gries, T., & Pilichowski, M. (2012). Stay in school or start working? The human capital investment decision under uncertainty and irreversibility. *Labour Economics*, 19(5), 706-717. doi: 10.1016/j.labeco.2012.04.005
- Bound, J., Lovenheim, M. F., & Turner, S. (2006). Increasing time to baccalaureate degree in the united states. *Education Finance and Policy*, 7, 375-424. doi: 10.1162/EDFP_a_00074
- Bourdieu, P. (1986). The forms of capital. In J. G. Richardson (Ed.), *Handbook of theory and research for the sociology of education* (pp. 241–258). Westport, CT: Greenwood Press Inc.
- Bourdieu, P., & Passeron, J. C. (1977). *Reproduction in education, society and culture*. Thousand Oaks, CA: Sage Publications.
- Braunstein, A., McGrath, M., & Pescatrice, D. (2000). Measuring the impact of financial factors on college persistence. *Journal of College Student Retention: Research, Theory & Practice*, 2(3), 191-203. doi: 10.2190/0TTM-U8RA-V8FX-FYVA
- Breen, R., & Goldthorpe, J. H. (1997). Explaining educational differentials towards a formal rational action theory. *Rationality and Society*, 9(3), 275-305. doi: 10.1177/104346397009003002
- Brown, S., Ortiz, A., & Taylor, K. (2006). *Educational attainment and risk preference*. Working Paper. Department of Economics, University of Sheffield. Retrieved from <http://eprints.whiterose.ac.uk/id/eprint/9938>
- Bryant, W. K., & Zick, C. D. (2006). *The economic organization of the household*. New York, NY: Cambridge University Press.
- Bureau of Labor Statistics. (2013a, September 14, 2006). National longitudinal surveys. from <http://www.bls.gov/nls/y97summary.htm>
- Bureau of Labor Statistics. (2013b). NLS news. Retrieved from <http://www.bls.gov/opub/nls/pdf/nls150.pdf>
- Bureau of Labor Statistics. (2013c). The NLSY97 sample: An introduction. Retrieved July 24, 2013, from <https://www.nlsinfo.org/content/cohorts/nlsy97/intro-to-the-sample/nlsy97-sample-introduction-0>
- Bureau of Labor Statistics. (n.d.-a). Administration of the CAT-ASVAB. *National Longitudinal Survey of Youth 1997*. Retrieved March 26, 2014, from <https://www.nlsinfo.org/content/cohorts/nlsy97/topical-guide/education/administration-cat-asvab-0>

- Bureau of Labor Statistics. (n.d.-b). Income, assets, and program participation. *National Longitudinal Survey of Youth 1997*. Retrieved March 29, 2014, from <https://www.nlsinfo.org/content/cohorts/nlsy97/topical-guide/income/assets-debts>
- Bureau of Labor Statistics. (n.d.-c). NLS investigator. Retrieved March 29, 2014, from <https://www.nlsinfo.org/investigator/pages/login.jsp>
- Bureau of Labor Statistics. (n.d.-d). Sample weights and design effects. *National Longitudinal Survey of Youth 1997*. Retrieved March 30, 2014, from <https://www.nlsinfo.org/content/cohorts/nlsy97/using-and-understanding-the-data/sample-weights-design-effects>
- Cabrera, A. F., & Castaneda, M. B. (1992). The convergence between two theories of college persistence. *Journal of Higher Education*, 63(2), 143-164. doi: 10.2307/1982157
- Callender, C., & Jackson, J. (2005). Does the fear of debt deter students from higher education? *Journal of Social Policy*, 34(4), 509. doi: 10.1017/S004727940500913X
- Callender, C., & Jackson, J. (2008). Does the fear of debt constrain choice of university and subject of study? *Studies in Higher Education*, 33(4), 405-429. doi: 10.1080/03075070802211802
- Cameron, A. C., & Trivedi, P. K. (2005). *Microeconometrics: Methods and applications*. New York, NY: Cambridge University Press.
- Card, D. (1999). The causal effect of education on earnings. *Handbook of Labor Economics*, 3, 1801-1863.
- Card, D., & Lemieux, T. (2001). Dropout and enrollment trends in the postwar period: What went wrong in the 1970s? *Risky behavior among youths: An economic analysis* (pp. 439-482): University of Chicago Press.
- Carnevale, A. P., Cheah, B., & Strohl, J. (2013). *Not all college degrees are created equal*. College, Majors, Unemployment, and Earnings. Georgetown University: Center on Education and the Workforce. Retrieved from <http://www9.georgetown.edu/grad/gppi/hpi/cew/pdfs/Unemployment.Final.update1.pdf>
- Catsiapis, G. (1987). A model of educational investment decisions. *The Review of Economics and Statistics*, 69(1), 33-41. doi: 10.2307/1937898
- Cha, K.-W., Weagley, R. O., & Reynolds, L. (2005). Parental borrowing for dependent children's higher education. *Journal of Family and Economic Issues*, 26(3), 299-321. doi: 10.1007/s10834-005-5900-y

- Chatterjee, S., & Ionescu, F. (2012). Insuring student loans against the financial risk of failing to complete college. *Quantitative Economics*, 3(3), 393-420. doi: 10.3982/qe100
- Chavas, J.-P. (2004). *Risk analysis in theory and practice*. San Diego, CA: Elsevier Academic Press.
- Checchi, D. (2006). *The economics of education: Human capital, family background and inequality*. New York, NY: Cambridge University Press.
- Cheng, X. (2013). An empirical test of the risk-return trade-off in educational choice. *Review of Business Research*, 13(1).
- Chiapa, C., Garrido, J. L., & Prina, S. (2012). The effect of social programs and exposure to professionals on the educational aspirations of the poor. *Economics of Education Review*, 31(5), 778-798. doi: 10.1016/j.econedurev.2012.05.006
- Christensen, S., Melder, J., & Weisbrod, B. A. (1975). Factors affecting college attendance. *The Journal of Human Resources*, 10(2), 174-188. doi: 10.2307/144825
- Christiansen, C., Joensen, J. S., & Nielsen, H. S. (2007). The risk-return trade-off in human capital investment. *Labour Economics*, 14(6), 971-986. doi: 10.1016/j.labeco.2007.06.001
- Clark, K. (2009). How much money should I borrow for college? *U.S. News & World Report*. <http://www.usnews.com/education/best-colleges/paying-for-college/articles/2009/06/09/how-much-money-should-i-borrow-for-college>
- Coleman, J. S. (1988). Social capital in the creation of human capital. *American Journal of Sociology*, 94(1), 95-120.
- College Affordability and Transparency Center. (2014). College scorecard. Retrieved February 4, 2014, from <http://www.whitehouse.gov/issues/education/higher-education/college-score-card>
- DeGroot, M. H., & Schervish, M. J. (2012). *Probability and statistics*. Boston, MA: Addison Wesley.
- Delaney, L., Harmon, C., & Redmond, C. (2011). Parental education, grade attainment and earnings expectations among university students. *Economics of Education Review*, 30(6), 1136-1152. doi: 10.1016/j.econedurev.2011.04.004
- Dickson, M., & Harmon, C. (2011). Economic returns to education: What we know, what we don't know, and where we are going-some brief pointers. *Economics of Education Review*, 30(6), 1118-1122. doi: 10.1016/j.econedurev.2011.08.003

- DiPrete, T. A., & Buchmann, C. (2013). *The rise of women: The growing gender gap in education and what it means for american schools*. New York, NY: Russell Sage Foundation.
- Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., & Wagner, G. G. (2011). Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the European Economic Association*, 9(3), 522-550. doi: 10.1111/j.1542-4774.2011.01015.x
- Dougherty, C. (2005). Why are the returns to schooling higher for women than for men? *Journal of Human Resources*, 40(4), 969-988.
- Dowd, A. C. (2004). Income and financial aid effects on persistence and degree attainment in public colleges. *Education Policy Analysis Archives*, 12(21).
- Dwyer, R. E., Hodson, R., & McCloud, L. (2013). Gender, debt, and dropping out of college. *Gender & Society*, 27(1), 30-55. doi: 10.1177/0891243212464906
- Dwyer, R. E., McCloud, L., & Hodson, R. (2012). Debt and graduation from American universities. *Social Forces*, 90(4), 1133-1155. doi: 10.1093/sf/sos072
- Dynarski, S. M. (2003). Does aid matter? Measuring the effect of student aid on college attendance and completion. *The American Economic Review*, 93(1).
- Dynarski, S. M. (2004). The new merit aid. In C. M. Hoxby (Ed.), *College choices: The economics of where to go, when to go, and how to pay for it* (pp. 63-100). Chicago, IL: University of Chicago Press.
- Dynarski, S. M. (2008). Building the stock of college-educated labor. *Journal of Human Resources*, 43(3), 576-610.
- Eaton, J., & Rosen, H. S. (1980). Taxation, human capital, and uncertainty. *The American Economic Review*, 70(4), 705-715.
- Fain, P. (2014). Gainful employment's partial unveiling. *Inside Higher Ed*. <http://www.insidehighered.com/news/2014/03/14/details-gainful-employment-proposal-expected-friday#sthash.jYTpNHZw.dpbs>
- Fang, H. (2006). Disentangling the college wage premium: Estimating a model with endogenous education choices. *International Economic Review*, 47(4), 1151-1185. doi: 10.1111/j.1468-2354.2006.00409.x
- Federal Reserve Bank of New York. (2012a). May 2012. *Quarterly Report on Household Debt and Credit*. from http://www.newyorkfed.org/research/national_economy/householdcredit/District_Report_Q12012.pdf

- Federal Reserve Bank of New York. (2012b). November 2012. *Quarterly Report on Household Debt and Credit*. Retrieved from http://www.newyorkfed.org/research/national_economy/householdcredit/District_Report_Q32012.pdf
- Field, A., & Miles, J. (2010). *Discovering statistics using SAS*. Thousand Oaks, CA: SAGE.
- Fishbein, M., & Ajzen, I. (1975). *Belief, attitude, intention, and behavior: An introduction to theory and research*. Reading, MA: Addison-Wesley Publishing Company.
- Fossen, F. M., & Glocker, D. (2011). Expected future earnings, taxation, and university enrollment. *International Tax and Public Finance*, 18(6), 688-723. doi: 10.1007/s10797-011-9173-0
- Fry, R. (2012). A record one-in-five households now owe student loan debt: Burden greatest on young, poor. *Social and Demographic Trends*. Washington, D.C.: Pew Research Center.
- Goldthorpe, J. H. (1996). Class analysis and the reorientation of class theory: The case of persisting differentials in educational attainment. *British Journal of Sociology*, 47(3), 481-505.
- Goyette, K. A. (2008). College for some to college for all: Social background, occupational expectations, and educational expectations over time. *Social Science Research*, 37(2), 461-484. doi: 10.1016/j.ssresearch.2008.02.002
- Grable, J. E. (2000). Financial risk tolerance and additional factors that affect risk taking in everyday money matters. *Journal of Business and Psychology*, 14(4), 625-630. doi: 10.1023/A:1022994314982
- Grable, J. E., & Lytton, R. H. (1998). Investor risk tolerance: Testing the efficacy of demographics as differentiating and classifying factors. *Financial Counseling and Planning*, 9, 1.
- Greene, W. H. (2012). *Econometric analysis* (7th ed.). Boston, MA: Prentice Hall.
- Hamilton, W. (2013). Student-loan delinquency rate hits danger zone, report says. *Los Angeles Times*. Retrieved from <http://www.latimes.com/business/money/la-fi-mo-student-loan-delinquency-20130129,0,548490.story>
- Hanna, S. D., Gutter, M. S., & Fan, J. X. (2001). A measure of risk tolerance based on economic theory. *Financial Counseling and Planning*, 12(2), 53-60.

- Hardekopf, B. (2013). More than half of student loans are now delinquent. *Forbes*. from <http://www.forbes.com/sites/moneybuilder/2013/02/01/alarming-number-of-student-loans-are-delinquent/>
- Harrast, S. A. (2004). Undergraduate borrowing: A study of debtor students and their ability to retire undergraduate loans. *Journal of Student Financial Aid*, 34(1), 21-37.
- Harris, M. (2011). Bad education. *n+1*. Retrieved from n+1 website: <http://nplusonemag.com/bad-education>
- Hartog, J., & Diaz-Serrano, L. (2007). Earnings risk and demand for higher education: A cross-section test for Spain. *Journal of Applied Economics*, 10(1), 1-28.
- Hastie, R., & Dawes, R. M. (2010). *Rational choice in an uncertain world: The psychology of judgment and decision making*. Thousand Oaks, CA: SAGE Publications.
- Heckman, J. J., Lochner, L. J., & Todd, P. E. (2008). Earnings functions and rates of return. *Journal of Human Capital*, 2(1), 1-31. doi: 10.1086/587037
- Heckman, S. J., Lim, H., & Montalto, C. P. (2013). Factors related to financial stress among college students. *Consumer Interests Annual*, 59.
- Hemelt, S. W., & Marcotte, D. E. (2011). The impact of tuition increases on enrollment at public colleges and universities. *Educational Evaluation and Policy Analysis*, 33(4), 435-457. doi: 10.3102/0162373711415261
- Hira, T. K., Anderson, M. M., & Petersen, K. (2000). Students' perceptions of their education debt and its impact on life after graduation. *Journal of Student Financial Aid*, 30(1), 7-19.
- Hogan, V., & Walker, I. (2007). Education choice under uncertainty: Implications for public policy. *Labour Economics*, 14(6), 894-912. doi: 10.1016/j.labeco.2007.06.009
- Hossler, D., Braxton, J., & Coopersmith, G. (Eds.). (1989). *Understanding student college choice* (Vol. 5). New York, NY: Agathon Press.
- Hossler, D., & Gallagher, K. S. (1987). Studying student college choice: A three-phase model and the implications for policymakers. *College and University*, 62(3), 207-221.
- Ishitani, T. T., & DesJardins, S. L. (2002). A longitudinal investigation of dropout from college in the United States. *Journal of College Student Retention*, 4(2), 173-201. doi: 10.2190/V4EN-NW42-742Q-2NTL

- Jacob, B., McCall, B., & Stange, K. (2013). *The consumption value of education: Implications for the postsecondary market*. Paper presented at the American Economic Association, San Diego, CA.
<http://www.aeaweb.org/aea/2013conference/program/retrieve.php?pdfid=469>
- Jacobs, B. (2007). Real options and human capital investment. *Labour Economics*, 14(6), 913-925. doi: 10.1016/j.labeco.2007.06.008
- Jaeger, D. A., & Page, M. E. (1996). Degrees matter: New evidence on sheepskin effects in the returns to education. *The Review of Economics and Statistics*, 78(4), 733-740. doi: 10.2307/2109960
- Kahneman, D. (2011). *Thinking, fast and slow*. New York, NY: Farrar, Straus and Giroux.
- Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2), 263-291. doi: 10.2307/1914185
- Kamenetz, A. (2006). *Generation debt: Why now is a terrible time to be young*. New York, NY: Riverhead Books.
- Kantrowitz, M. (2013). Calculator - how much to borrow. Retrieved May 5, 2013 from <http://www.finaid.org/calculators/howmuchtoborrow.phtml>
- Kmett, C. M., Arkes, H. R., & Jones, S. K. (1999). The influence of decision aids on high school students' satisfaction with their college choice decision. *Personality and Social Psychology Bulletin*, 25(10), 1293-1301. doi: 10.1177/0146167299258010
- Knight, F. H. (1921). *Risk, uncertainty and profit*. Boston, MA: Houghton Mifflin Company.
- Kodde, D. A. (1986). Uncertainty and the demand for education. *The Review of Economics and Statistics*, 68(3), 460-467. doi: 10.2307/1926023
- Koehler, D. J., & Harvey, N. (Eds.). (2008). *Blackwell handbook of judgment and decision making*. Malden, MA: Blackwell Publishing.
- Koerselman, K., & Uusitalo, R. (2013). *The risk and return of human capital investments*. IZA Discussion Papers. Institute for the Study of Labor. Retrieved from <http://hdl.handle.net/10419/89880>
- Kreps, D. M. (1990). *A course in microeconomic theory*. Princeton, NJ: Princeton University Press.
- Kulatilaka, N. (1995). The value of flexibility: A general model of real options. In L. Trigeorgis (Ed.), *Real options in capital investments*. Westport, CT: Praeger Publishers.

- Laitinen, A. (2012). *Cracking the credit hour*. New America Foundation and Education Sector. Retrieved from http://www.educationsector.org/sites/default/files/publications/Cracking_the_Credit_Hour_0.pdf
- Lassila, N. E. (2011). Effects of tuition price, grant aid, and institutional revenue on low-income student enrollment. *Journal of Student Financial Aid*, 41(3), 24-41.
- Letkiewicz, J. L., Lim, H., Heckman, S. J., Bartholomae, S., Fox, J. J., & Montalto, C. P. (in press). The path to graduation: Factors predicting on-time graduation rates. *Journal of College Student Retention: Research, Theory & Practice*
- Levhari, D., & Weiss, Y. (1974). The effect of risk on the investment in human capital. *The American Economic Review*, 64(6), 950-963.
- Li, D. (2008). Degree attainment of undergraduate student borrowers in four-year institutions: A multilevel analysis. *Journal of Student Financial Aid*, 37(3), 5-16.
- Light, A. (2001). In-school work experience and the returns to schooling. *Journal of Labor Economics*, 19(1), 65-93. doi: 10.1086/209980
- Light, A., & Ahn, T. (2010). Divorce as risky behavior. *Demography*, 47(4), 895-921. doi: 10.1007/BF03213732
- Lillard, D., & Gerner, J. (1999). Getting to the ivy league: How family composition affects college choice. *The Journal of Higher Education*, 70(6), 706-730. doi: 10.2307/2649172
- Lovenheim, M. F., & Reynolds, C. L. (2011). Changes in postsecondary choices by ability and income: Evidence from the national longitudinal surveys of youth. *Journal of Human Capital*, 5(1), 70-109. doi: 10.1086/660123
- Mas-Colell, A., Whinston, M. D., & Green, J. R. (1995). *Microeconomic theory*. New York, NY: Oxford University Press.
- Mazza, J. (2011). *Does risk matter? A semiparametric model for educational choices in the presence of uncertainty*. Working Papers. Social Science Research Network. Retrieved from <http://ssrn.com/abstract=2016104>
- Mazza, J., van Ophem, H., & Hartog, J. (2013). Unobserved heterogeneity and risk in wage variance: Does more schooling reduce earnings risk? *Labour Economics*, 24, 323-338. doi: 10.1016/j.labeco.2013.09.007
- Mincer, J. (1958). Investment in human capital and personal income distribution. *The Journal of Political Economy*, 66(4), 281-302.

- Mincer, J. (1974). *Schooling, experience, and earnings*. New York, NY: National Bureau of Economic Research, Columbia University Press.
- National Center for Education Statistics. (2013a). Fast facts: Immediate transition to college. Retrieved February 14, 2014, from <https://nces.ed.gov/fastfacts/display.asp?id=51>
- National Center for Education Statistics. (2013b). Surveys and programs. Retrieved September 23, 2013, from <http://nces.ed.gov/surveys/SurveyGroups.asp?Group=2>
- Neelakantan, U. (2010). Estimation and impact of gender differences in risk tolerance. *Economic Inquiry*, 48(1), 228-233. doi: 10.1111/j.1465-7295.2009.00251.x
- Oreopoulos, P., & Petronijevic, U. (2013). *Making college worth it: A review of research on the returns to higher education*. NBER Working Paper. National Bureau of Economic Research. Retrieved from <http://www.nber.org/papers/w19053>
- Oreopoulos, P., & Salvanes, K. G. (2011). Priceless: The nonpecuniary benefits of schooling. *Journal of Economic Perspectives*, 25(1), 159-184. doi: 10.1257/jep.25.1.159
- Palacios-Huerta, I. (2003). An empirical analysis of the risk properties of human capital returns. *The American Economic Review*, 948-964.
- Paulsen, M. B. (1990). *College choice: Understanding student enrollment behavior*. ASHE-ERIC Higher Education Report No. 6. Association for the Study of Higher Education.
- Paulsen, M. B., & St. John, E. P. (1997). The financial nexus between college choice and persistence. *New Directions for Institutional Research*, 1997(95), 65-82. doi: 10.1002/ir.9504
- Paulsen, M. B., & St. John, E. P. (2002). Social class and college costs: Examining the financial nexus between college choice and persistence. *Journal of Higher Education*, 73(2), 189-236. doi: 10.1353/jhe.2002.0023
- Perna, L. W. (1998). The contribution of financial aid to undergraduate persistence. *Journal of Student Financial Aid*, 28(3), 25-40.
- Perna, L. W. (2000). Differences in the decision to attend college among african americans, hispanics, and whites. *Journal of Higher Education*, 71(2), 117-141. doi: 10.2307/2649245
- Perna, L. W. (2006). Studying college access and choice: A proposed conceptual model. In J. C. Smart (Ed.), *Higher education: Handbook of theory and research* (Vol. 21, pp. 99-157). Netherlands: Springer.

- Perna, L. W. (2008). Understanding high school students' willingness to borrow to pay college prices. *Research in Higher Education*, 49(7), 589-606. doi: 10.1007/s11162-008-9095-6
- Perna, L. W., & Titus, M. A. (2004). Understanding differences in the choice of college attended: The role of state public policies. *Review of Higher Education*, 27(4), 501-525. doi: 10.1353/rhe.2004.0020
- Pratt, J. W. (1964). Risk aversion in the small and in the large. *Econometrica*, 32(1/2), 122-136. doi: 10.2307/1913738
- Psacharopoulos, G. (1994). Returns to investment in education: A global update. *World Development*, 22(9), 1325-1343. doi: 10.1016/0305-750X(94)90007-8
- Psacharopoulos, G., & Patrinos, H. A. (2004). Returns to investment in education: A further update. *Education Economics*, 12(2), 111-134. doi: 10.1080/0964529042000239140
- Putnam, R. D. (1995). Bowling alone: America's declining social capital. *Journal of Democracy*, 6(1), 65-78. doi: 10.1353/jod.1995.0002
- Putnam, R. D. (2001). *Bowling alone: The collapse and revival of american community*. New York, NY: Simon & Schuster.
- Reisel, L. (2013). Is more always better? Early career returns to education in the united states and norway. *Research in Social Stratification and Mobility*, 31(0), 49-68. doi: 10.1016/j.rssm.2012.10.002
- Robb, C. A., Moody, B., & Abdel-Ghany, M. (2011). College student persistence to degree: The burden of debt. *Journal of College Student Retention: Research, Theory and Practice*, 13(4), 431-456. doi: 10.2190/CS.13.4.b
- Roksa, J. (2010). Bachelor's degree completion across state contexts: Does the distribution of enrollments make a difference? *Research in Higher Education*, 51(1), 1-20. doi: 10.1007/s11162-009-9146-7
- Sahm, C. R. (2012). How much does risk tolerance change? *Quarterly Journal of Finance*, 02(04), 1250020. doi: 10.1142/S2010139212500206
- SAS Institute. (2014). Binomial proportions. *SAS/STAT(R) 9.22 User's Guide*. Retrieved May 28, 2014, from https://support.sas.com/documentation/cdl/en/statug/63347/HTML/default/viewer.htm#statug_freq_a0000000660.htm
- Savoca, E. (1990). Another look at the demand for higher education: Measuring the price sensitivity of the decision to apply to college. *Economics of Education Review*, 9(2), 123-134. doi: 10.1016/0272-7757(90)90040-C

- Scharff, R. L., & Viscusi, W. K. (2011). Heterogeneous rates of time preference and the decision to smoke. *Economic Inquiry*, 49(4), 959-972. doi: 10.1111/j.1465-7295.2009.00191.x
- Schneider, B. L., & Stevenson, D. (2000). *The ambitious generation: Americas teenagers, motivated but directionless*. New Haven, CT: Yale University Press.
- Selingo, J. (2014). To reach the new market for education, colleges have some learning to do. *The Chronicle of Higher Education*. <http://chronicle.com/article/To-Reach-the-New-Market-for/145499/>
- Spence, M. (1973). Job market signaling. *The Quarterly Journal of Economics*, 87(3), 355-374. doi: 10.2307/1882010
- St. John, E. P., Paulsen, M. B., & Carter, D. F. (2005). Diversity, college costs, and postsecondary opportunity: An examination of the financial nexus between college choice and persistence for african americans and whites. *Journal of Higher Education*, 76(5), 545-569. doi: 10.1353/jhe.2005.0035
- St. John, E. P., Paulsen, M. B., & Starkey, J. B. (1996). The nexus between college choice and persistence. *Research in Higher Education*, 37(2), 175-220. doi: 10.2307/40196172
- STATA. (2014). Stata 13 user manuals: Margins. Retrieved March 28, 2014, 2014, from <http://www.stata.com/manuals13/rmargins.pdf>
- Sung, J., & Hanna, S. D. (1996). Factors related to risk tolerance. *Financial Counseling and Planning*, 7(1), 11-20.
- Thaler, R. H., & Sunstein, C. R. (2008). *Nudge: Improving decisions about health, wealth, and happiness*. New Haven, CT: Yale University Press.
- The U.T. System Student Debt Reduction Task Force. (2012). *College "credit": Reducing unmanageable student debt and maximizing return on education*. University of Texas System. Retrieved from <http://www.utsystem.edu/sites/utsfiles/news/assets/college-credit-student-debt-report.pdf>
- Tinto, V. (1975). Dropout from higher education: A theoretical synthesis of recent research. *Review of Educational Research*, 45(1), 89-125. doi: 10.3102/00346543045001089
- Titus, M. A. (2006a). No college student left behind: The influence of financial aspects of a state's higher education policy on college completion. *Review of Higher Education*, 29(3), 293-317. doi: 10.1353/rhe.2006.0018

- Titus, M. A. (2006b). Understanding college degree completion of students with low socioeconomic status: The influence of the institutional financial context. *Research in Higher Education*, 47(4), 371-398. doi: 10.1007/s11162-005-9000-5
- Torres, V., Gross, J. P. K., & Dadashova, A. (2010). Traditional-age students becoming at-risk: Does working threaten college students' academic success? *Journal of College Student Retention: Research, Theory and Practice*, 12(1), 51-68. doi: 10.2190/CS.12.1.d
- Trombitas, K. S. (2012). *Financial stress: An everyday reality for college students*. White Paper. Inceptia. Lincoln, NE. Retrieved from https://www.inceptia.org/PDF/Inceptia_FinancialStress_whitepaper.pdf
- Turner, S. (2004). Going to college and finishing college. Explaining different educational outcomes. In C. M. Hoxby (Ed.), *College choices: The economics of where to go, when to go, and how to pay for it* (pp. 13-62): University of Chicago Press.
- U.S. Census Bureau. (n.d.). Census regions and divisions of the United States. Retrieved April 22, 2014, from http://www.census.gov/geo/maps-data/maps/pdfs/reference/us_regdiv.pdf
- U.S. Department of Transportation. (n.d.). Census regions of the United States. Retrieved April 22, 2014, from http://www.rita.dot.gov/bts/sites/rita.dot.gov.bts/files/publications/america_on_the_go/us_business_travel/html/figure_02.html
- Varian, H. (1992). *Microeconomic analysis* (3rd ed.). New York, NY: W.W. Norton & Company.
- von Neumann, J., & Morgenstern, O. (1944). *Theory of games and economic behavior*. Princeton, NJ: Princeton University Press.
- Wang, X., & Wickersham, K. (2013). Postsecondary co-enrollment and baccalaureate completion: A look at both beginning 4-year college students and baccalaureate aspirants beginning at community colleges. *Research in Higher Education*, 1-30. doi: 10.1007/s11162-013-9317-4
- Weber, E. U., Blais, A.-R., & Betz, N. E. (2002). A domain-specific risk-attitude scale: Measuring risk perceptions and risk behaviors. *Journal of Behavioral Decision Making*, 15(4), 263-290. doi: 10.1002/bdm.414
- Williams, J. T. (1979). Uncertainty and the accumulation of human capital over the life cycle. *The Journal of Business*, 52(4), 521-548. doi: 10.2307/2352444

- Winston, G., & Zimmerman, D. (2004). Peer effects in higher education. In C. M. Hoxby (Ed.), *College choices: The economics of where to go, when to go, and how to pay for it* (pp. 395-424). Chicago, IL: University of Chicago Press.
- Wooldridge, J. M. (2009). *Introductory econometrics: A modern approach*. Mason, OH: South-Western Cengage Learning.
- Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. Cambridge, MA: MIT Press.
- Zhan, M. (2012). *The impact of youth debt on college graduation*. Assets and Education Research Symposium. Retrieved from <http://csd.wustl.edu/Publications/Documents/WP12-11.pdf>

Appendix A: Age-Adjusted Risk Tolerance Computation

The estimation of the age-adjusted risk tolerance measure followed the procedure and methodology outlined by Light and Ahn (2010). This discussion is organized into three sections. The first section outlines the initial step of this procedure which involves computing the upper and lower bounds for the coefficient of relative risk tolerance from the income gamble questions. The second section reviews Light and Ahn's estimation procedure and the third section gives an overview of the computation in the current study that used Light and Ahn's parameter estimates.

A.1 Upper and Lower Bounds of the Coefficient of Relative Risk Tolerance

By applying expected utility theory, the upper and lower bounds of the Arrow-Pratt coefficient of relative risk tolerance can be derived from the income gamble question in Section 5.3.2 (e.g., see Barsky et al., 1997; Light & Ahn, 2010). As an example, consider the third risk category. Respondents in this category chose to accept the gamble that their income may double or be reduced by one third in the first question and then rejected the gamble that income may double or be reduced by one half. In expected utility terms, the first choice implies (A.1) and the second choice implies (A.2):

$$U(Y) \leq 0.5U(2Y) + 0.5U\left(\frac{2}{3}Y\right), \quad (\text{A.1})$$

$$U(Y) > 0.5U(2Y) + 0.5U\left(\frac{1}{2}Y\right). \quad (\text{A.2})$$

(A.1) and (A.2) imply the following:

$$0.5U(2Y) + 0.5U\left(\frac{1}{2}Y\right) < U(Y) \leq 0.5U(2Y) + 0.5U\left(\frac{2}{3}Y\right). \quad (\text{A.3})$$

Assuming a utility function that exhibits constant relative risk aversion, such as,

$$U(C_i) = \frac{C_i^{1-1/\rho_{it}}}{1-1/\rho_{it}}, \quad (\text{A.4})$$

the upper and lower bounds can be computed by substituting (A.3) into (A.4); that is,

$$0.5 \left(\frac{2Y^{1-1/\rho_{it}}}{1-1/\rho_{it}} \right) + 0.5 \left(\frac{\frac{1}{2}Y^{1-1/\rho_{it}}}{1-1/\rho_{it}} \right) < \left(\frac{Y^{1-1/\rho_{it}}}{1-1/\rho_{it}} \right) \leq 0.5 \left(\frac{2Y^{1-1/\rho_{it}}}{1-1/\rho_{it}} \right) + 0.5 \left(\frac{\frac{2}{3}Y^{1-1/\rho_{it}}}{1-1/\rho_{it}} \right). \quad (\text{A.5})$$

Simplifying (A.5) implies that the coefficient of relative risk tolerance of an individual in the third risk category must lie between 0.5 and 1.0. The preceding procedure can be used to infer the upper and lower bounds for the remaining categories, shown in Table A.1.

Table A.1. Upper and Lower Bounds for Risk Tolerance

Relative Risk Tolerance Category	Bounds for the Coefficient of Relative Risk Tolerance	
	Lower	Upper
1 (lowest)	0	0.27
2	0.27	0.5
3	0.5	1.0
4 (Highest)	1.0	∞

Source: Light and Ahn (2010) , Barsky et al. (1997)

A.2 Light and Ahn's Estimation Method

Light and Ahn used (A.6) to estimate the coefficient of relative risk tolerance, ρ , for individual i at time t :

$$\log \rho_{it} = \beta AGE_{it} + \alpha_i + u_{it}. \quad (\text{A.6})$$

(A.6) was estimated separately for men and women and the parameter estimates are shown in Table A.2.

Table A.2. Light and Ahn's (2010) Estimates of Risk Preference Parameters

Parameter	Women		Men	
	Estimate	SE	Estimate	SE
β	-0.035	0.003	-0.047	0.003
$\bar{\alpha}$	-0.178	0.104	0.525	0.114
σ_{α}	1.277	0.029	1.390	0.032
σ_u	1.131	0.019	1.222	0.021

Source: Light and Ahn (2010), p. 903.

Since the NLSY97 has only one round of collection on the income gamble question, (A.6) is modified to the following:

$$\log \rho_{it} = \beta AGE_{it} + \bar{\alpha} + \underbrace{(\alpha_i - \bar{\alpha}) + u_{it}}_{\varepsilon_{it}}. \quad (\text{A.7})$$

$$\text{where } \text{Var}(\varepsilon) = \sigma_{\alpha}^2 + \sigma_u^2.$$

The primary difference is that the average individual random effect needed to be used since the individual random effects could not be estimated in the NLSY97.

The predicted coefficient of relative risk tolerance at time τ given the response to the income gamble question at time t is modeled as follows:

$$E[\rho_{it}|c_{it} = j, AGE_{it}, AGE_{it}] = \exp\left(\hat{\alpha} + \hat{\beta}AGE_{it} + \frac{1}{2}\hat{\sigma}_{\alpha}^2\right) \frac{P(\log \rho_j < \log \rho_{it} + \hat{\sigma}_{\alpha}^2 < \log \bar{\rho}_j)}{P(\log \rho_j < \log \rho_{it} < \log \bar{\rho}_j)}. \quad (A.8)$$

where $c_{it} = j$ ($j = 1, 2, 3, 4$) is the risk category as described in Section

5.3.2.

By substituting (A.7) into (A.8),

$$E[\rho_{it}|c_{it} = j, AGE_{it}, AGE_{it}] = \exp\left(\hat{\alpha} + \hat{\beta}AGE_{it} + \frac{1}{2}\hat{\sigma}_{\alpha}^2\right) \frac{P(\log \rho_j - \hat{\beta}AGE_{it} - \bar{\alpha} - \hat{\sigma}_{\alpha}^2 < \varepsilon_{it} < \log \bar{\rho}_j - \hat{\beta}AGE_{it} - \bar{\alpha} - \hat{\sigma}_{\alpha}^2)}{P(\log \rho_j - \hat{\beta}AGE_{it} - \bar{\alpha} < \varepsilon_{it} < \log \bar{\rho}_j - \hat{\beta}AGE_{it} - \bar{\alpha})}. \quad (A.9)$$

Finally, by expanding (A.9), the full computation is shown in A.10.

$$E[\rho_{it}|c_{it} = j, AGE_{it}, AGE_{it}] = \exp\left(\hat{\alpha} + \hat{\beta}AGE_{it} + \frac{1}{2}\hat{\sigma}_{\alpha}^2\right) \frac{\Phi\left(\frac{\log \bar{\rho}_j - \hat{\beta}AGE_{it} - \bar{\alpha} - \hat{\sigma}_{\alpha}^2}{\sqrt{\hat{\sigma}_{\alpha}^2 + \hat{\sigma}_u^2}}\right) - \Phi\left(\frac{\log \rho_j - \hat{\beta}AGE_{it} - \bar{\alpha} - \hat{\sigma}_{\alpha}^2}{\sqrt{\hat{\sigma}_{\alpha}^2 + \hat{\sigma}_u^2}}\right)}{\Phi\left(\frac{\log \bar{\rho}_j - \hat{\beta}AGE_{it} - \bar{\alpha}}{\sqrt{\hat{\sigma}_{\alpha}^2 + \hat{\sigma}_u^2}}\right) - \Phi\left(\frac{\log \rho_j - \hat{\beta}AGE_{it} - \bar{\alpha}}{\sqrt{\hat{\sigma}_{\alpha}^2 + \hat{\sigma}_u^2}}\right)}, \quad (A.10)$$

where Φ is the standard normal distribution.

The following code was used in SAS 9.3 to compute the coefficient of relative risk tolerance using (A.10).

```

if female=1 then do;
    beta=-0.035;
    alphabar=-0.178;
    sigmaalpha=1.277;
    sigmau=1.131;
end;
if female=0 then do;
    beta=-0.047;
    alphabar=0.525;
    sigmaalpha=1.390;
    sigmau=1.222;
end;
* age when risk question was asked=riskdate*;
AgeRisk=(riskdate-DOB)/365;
sigalphasq=sigmaalpha*sigmaalpha;
sigusq=sigmau*sigmau;
if B SorCCap=0 then enrollage=(max(HSGraddate,GEDdate)-DOB)/365;
else if B SorCCap=1 then enrollage=min(ageenrollCC, ageenrollBS);

```

```

AgeADJ=exp(alphabar+(beta*enrollage)+ (.5*sigalphasq));

if RRACat=1 then do;
    NUp=( (log(.27)-(beta*agerisk)-alphabar-
sigalphasq)/(sqrt(sigalphasq+sigusq))));
    NDIFF=CDF('Normal',NUp);
    DUp=( (log(.27)-(beta*agerisk)-
alphabar)/(sqrt(sigalphasq+sigusq))));
    DLo=( (0-(beta*agerisk)-alphabar)/(sqrt(sigalphasq+sigusq))));
    DDIFF= CDF('Normal',DUp)-0;
end;
if RRACat=2 then do;
    NUp=( (log(.5)-(beta*agerisk)-alphabar-
sigalphasq)/(sqrt(sigalphasq+sigusq))));
    NLo=( (log(.27)-(beta*agerisk)-alphabar-
sigalphasq)/(sqrt(sigalphasq+sigusq))));
    NDIFF=CDF('Normal',NUp)-CDF('Normal',NLo);
    DUp=( (log(.5)-(beta*agerisk)-
alphabar)/(sqrt(sigalphasq+sigusq))));
    DLo=( (log(.27)-(beta*agerisk)-
alphabar)/(sqrt(sigalphasq+sigusq))));
    DDIFF=CDF('Normal',DUp)-CDF('Normal',DLo);
end;
if RRACat=3 then do;
    NUp=( (log(1)-(beta*agerisk)-alphabar-
sigalphasq)/(sqrt(sigalphasq+sigusq))));
    NLo=( (log(.5)-(beta*agerisk)-alphabar-
sigalphasq)/(sqrt(sigalphasq+sigusq))));
    NDIFF=CDF('Normal',NUp)-CDF('Normal',NLo);
    DUp=( (log(1)-(beta*agerisk)-
alphabar)/(sqrt(sigalphasq+sigusq))));
    DLo=( (log(.5)-(beta*agerisk)-
alphabar)/(sqrt(sigalphasq+sigusq))));
    DDIFF=CDF('Normal',DUp)-CDF('Normal',DLo);
end;
if RRACat=4 then do;
    NUp=( (1-(beta*agerisk)-alphabar-
sigalphasq)/(sqrt(sigalphasq+sigusq))));
    NLo=( (log(.5)-(beta*agerisk)-alphabar-
sigalphasq)/(sqrt(sigalphasq+sigusq))));
    NDIFF=1-CDF('Normal',NLo);
    DUp=( (1-(beta*agerisk)-alphabar)/(sqrt(sigalphasq+sigusq))));
    DLo=( (log(.5)-(beta*agerisk)-
alphabar)/(sqrt(sigalphasq+sigusq))));
    DDIFF=1-CDF('Normal',DLo);
end;
Prob=NDIFF/DDIFF;
RiskTol=AgeADJ*Prob;

```


Appendix B: Proportion and Means Tests in Table 6.1

This section outlines the three types of hypothesis testing that were used in Table 6.1, including each of the following: independent Samples t-Tests, test of homogeneity, and binomial proportion tests.

Independent Samples *t*-Test

The independent samples t-test evaluates whether the mean for one sample is equal to the mean of another sample. As described in Field and Miles (2010, pp. 280-283), if the sample sizes are not equal, researchers should use estimate of the pooled variance (s_p^2) in the calculation of the test statistic. Let n_i represent the number of observations in sample i and let s_i^2 represent the variance of sample i :

$$s_p^2 = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1 + n_2 - 2}. \quad (\text{B.1})$$

The t test statistic is computed by dividing the mean difference between the two samples ($\bar{X}_1 - \bar{X}_2$) by the standard error of the sampling distribution (Field & Miles, 2010, pp. 280-283):

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2}}}. \quad (\text{B.2})$$

The test statistic, t , follows a t-distribution with $(n_1 + n_2 - 2)$ degrees of freedom.

χ^2 test of Homogeneity

For a response variable of C levels, the χ^2 test of homogeneity tests whether or not the proportion of $c=1, \dots, c=C$ is the same across all levels of an independent variable with R levels (DeGroot & Schervish, 2012, pp. 648-649). Let i represent the i^{th} level of the independent variable and j represent the j^{th} level of the response variable. For an $i \times j$ table, the null hypothesis is that all of the proportions for a given level of j are the same across all levels of i .

$$H_0: p_{1j} = p_{2j} = \dots = p_{Rj}, \text{ for } j = 1, \dots, C. \quad (\text{B.3})$$

The test statistic weights the squared difference between the observed cell counts and the expected cell counts by the expected cell counts (DeGroot & Schervish, 2012, pp. 648-649). This weighted square difference is then summed over all of the cells to calculate the test statistic, Q . That is,

$$Q = \sum_{i=1}^R \sum_{j=1}^C \frac{(N_{ij} - \hat{E}_{ij})^2}{\hat{E}_{ij}}, \quad (\text{B.4})$$

$$\text{where} \quad \hat{E}_{ij} = \frac{N_{i+} N_{+j}}{n}.$$

Q follows an approximate χ^2 distribution with $(R - 1)(C - 1)$ degrees of freedom.

Binomial Proportion

The binomial proportion test is the standard z test of proportions that evaluates whether or not a sample proportion is equal to the proportion under null hypothesis, p_0 (SAS Institute, 2014). The sample proportion, $\hat{p}$, is calculated as the number of observations of the subsample of interested (n_1) divided by the total sample (n):

$$\hat{p} = n_1/n. \quad (\text{B.5})$$

The standard error of the sampling distribution is computed as:

$$se(\hat{p}) = \sqrt{\hat{p}(1 - \hat{p})/n}. \quad (\text{B.6})$$

The test statistic, z , follows the standard normal distribution $N \sim (0,1)$ and is computed by dividing the difference between the observed and hypothesized proportion by the standard error of the sampling distribution.

$$z = (\hat{p} - p_0)/se. \quad (\text{B.7})$$

Appendix C: Average Marginal Effects

As explained in Wooldridge (2010, pp. 566-567), for a continuous variable in a binary response model, the marginal effect is

$$\frac{\partial p(x)}{\partial x_j} = G(X'\beta)\beta_j, \quad (\text{C.1})$$

where G is the logistic function.

For a binary explanatory variable, x_K , the marginal effect from changing from zero to one is

$$G(\beta_0 + \beta_1 x_1 + \cdots + \beta_{K-1} x_{K-1} + \beta_K) - G(\beta_0 + \beta_1 x_1 + \cdots + \beta_{K-1} x_{K-1}). \quad (\text{C.2})$$

The average marginal effect computes the marginal effect for each observation in the sample and averages it across the entire sample. That is, for a continuous variable,

$$\beta_j \left[\frac{1}{N} \sum_{i=1}^N G(X'_i \beta) \right], \quad (\text{C.3})$$

and for a binary variable,

$$\beta_K \left[\frac{1}{N} \sum_{i=1}^N G(\beta_0 + \beta_1 x_1 + \cdots + \beta_{K-1} x_{K-1} + \beta_K) - G(\beta_0 + \beta_1 x_1 + \cdots + \beta_{K-1} x_{K-1}) \right]. \quad (\text{C.4})$$

For more information, see Wooldridge (2010, pp. 566-567) or the STATA manual regarding the margins command (STATA, 2014).