College Premiums and the Economic Value of an Undergraduate Degree

by

Nathaniel Campagne
Class of 2015

A thesis submitted to the faculty of Wesleyan University in partial fulfillment of the requirements for the Degree of Bachelor of Arts with Departmental Honors in Economics.

Middletown, Connecticut April, 2015
Acknowledgments

First and foremost, I want to thank Professor Richard Grossman for his advice and support, freely given and deeply appreciated throughout this process. I would also like to thank Professors Jacobsen, Khamis, Keats, and Kuenzel for providing incredibly helpful insight at numerous points along the way. For their ceaseless willingness to discuss this thesis with me, I want to express my gratitude to my parents and Andrew Hove. I’d also like to thank Leib Sutcher for convincing me to do a thesis in the first place, then convincing me to use Stata, and then teaching me how. Lastly, I want to thank Linda Mascaro for having the answer to every question there is.
Abstract

This thesis analyzes the difference in income between college and high school graduates—referred to as the “college premium”—between 1971 and 2014. Data from the Current Population Survey are used to investigate trends, demographic factors, and several determinants of college premiums. The results presented here indicate that premiums, after falling for the first half of the 1970s, increased dramatically over the remainder of the period under study. In addition, premiums differed significantly among certain demographic groups. Special consideration is given to the impact of selection bias and the popularity of different fields of study, in order to show that the economic value of a college degree increased with college premiums.
Introduction

Economists agree that income inequality has been rising in the United States since the 1980s, even if the forces underlying this increase are neither universally understood nor agreed upon (Gordon and Dew-Becker, 2008). One important observation regarding the distribution of wealth in the U.S. is that a large portion of the increase in income inequality can be explained by widening earnings differentials between people with different levels of education (Goldin and Katz, 2007b). Thus, an analysis of the trends and determinants of these differentials has the potential to deepen our understanding of the increasingly unequal distribution of wealth in the United States.

Perhaps the most frequently discussed education-related earnings differential is the one between college and high school graduates. This figure has many names in the literature—returns to higher education, wage premium, college differential, etc.—but in this thesis, I use the label “college premium” as an umbrella term for the mean percentage difference in income, wages, or earnings.
between college and high school graduates. The primary objectives of this thesis are to analyze college premiums over the last 44 years and explore differences in college premiums between certain demographic groups. A secondary objective will be to investigate several factors that may be driving the trends in college premiums, with the ultimate goal of applying the results to questions of income inequality and human capital accumulation in the United States.

This thesis is by no means the first attempt at analyzing college premiums. Since the 1970s, a large body of literature has been published on the topic, much of which is discussed in subsequent chapters of this thesis. I hope to enrich the existing literature in the following ways: (1) I will add evidence aimed at advancing several of the ongoing debates related to college premiums, (2) I will improve upon the methodologies used in previous studies in order to present a more accurate and precise analysis of college premiums, and (3) I will explore the question of what can and cannot be inferred from the trends in college
premiums, in order to connect my findings to several other areas of research.

This thesis is organized into four chapters. Chapter 1 summarizes and discusses the existing literature on college premiums. Chapter 2 presents the first part of my analysis, focused on estimating college premiums over time and across demographic groups. Chapter 3 is the second part of my analysis, which seeks to refine and expand the results from Chapter 2 and discuss their implications. Chapter 4 concludes, summarizes the overarching findings and takeaways, and outlines potential directions for future research.
Overview of the College Premium

Basic Theories

Social scientists have long agreed that pursuing postsecondary education in the United States yields major, long-lasting economic benefits (Paulsen, 1998; Fang, 2006; Diprete and Buchmann, 2006). From an economic perspective, going to college would be a highly irrational decision if not for the increased earnings expected from completing a college degree. There are two leading explanations for why possessing a college degree leads to improved economic outcomes. The first, generally referred to as productivity enhancement, is related to the acquiring of useful knowledge and skills. This explanation maintains that the experience of obtaining a bachelor’s degree makes the graduate a more valuable asset in the workplace, so that employers will bid up the price of their labor and they will be paid more than a ceteris paribus individual without an undergraduate education. Although many economists support this explanation, others believe there is a
second crucial mechanism at play, which they refer to as ability signaling (Fang, 2006). The premise of this theory is that degree holders tend to be naturally more valuable in the workplace than people without a degree, since getting into and graduating from college requires a certain degree of intellectual competence. Thus, the process of obtaining a postsecondary education is a way of signaling to future employers that the prospective employee is valuable.

Both explanations have important implications for the questions with which this thesis is concerned. The productivity enhancement theory, favored explicitly by Fang (2006) and implicitly by Riley (1982), suggests that the benefit of pursuing a college degree lies in the process of obtaining it. This perspective highlights the economic value of higher education itself, emphasizing the value of the educative process. The ability signaling hypothesis also has ample support, including papers by Arcidiacono, Bayer, and Gizmo (2010), and Oreopoulos and Petronijevic (2013). In contrast to productivity enhancement, ability signaling suggests that the primary economic value of college is
simply to make the degree holder stand out to employers. In doing so, it implies that the college-educated cohort would still be more valuable in the workplace—and thus entitled to higher earnings—even if they never went to college. This selection effect, referred to in this thesis as either selection bias or ability bias, theoretically prevents one from estimating the economic value of higher education by simply looking at differences in outcomes between people who possess a college degree and people who do not. It is my belief that the ability signaling explanation is valid, and that a portion of these differences does indeed stem from selection bias. For that reason, the term “college premium” is used to describe college-high school differentials in wages, earnings, or income. Terms like “the economic value of college” and “returns to higher education” are reserved for the college premium minus the magnitude of the ability bias.

It should be noted that the productivity enhancement theory does not rule out the existence of a selection bias; in fact, such a bias is still plausible even if one ignores the ability signaling hypothesis. The difference is that the ability signaling theory
directly implies a selection bias, whereas the productivity enhancement hypothesis takes no stand on whether or not this bias exists.

Positive Selection Versus Negative Selection

In addition to the theoretical framework discussed above, economists have also investigated whether or not college premiums vary among different demographic groups. There are two fundamental hypotheses with respect to this line of research, neatly outlined by Brand and Xie (2010). On one hand, there is the positive selection hypothesis, which asserts that the demographic groups that are most likely to attend college also benefit the most from it. This hypothesis stems from the assumption of rational agents, arguing that if going to college is more common for a given demographic group, it must be because members of that group benefit more from owning a degree, and are thus more likely to choose to attend college.

This theory is opposed by the negative selection hypothesis, which contends that the demographic groups least likely to go to
college receive the largest benefit from doing so. Rooted in sociological studies of motivation and justification, this theory is favored by a number of economists (Brand and Xie, 2010). The concept is that demographic groups that are more likely to attend college—for example, the white upper class—typically have less career-oriented goals while pursuing their degree. They may be more likely to choose programs based solely on interest and enroll in liberal arts programs or other majors associated with lower earnings after graduation. In contrast, demographic groups that have lower college enrollment rates, such as the black lower- and middle-classes, tend to choose degree programs that have higher average payoffs. These groups, according to the theory, are more likely to view college as a vehicle for upward mobility than as a purely intellectual pursuit, choosing their schools and majors with the primary goal of maximizing future income. For example, they may be more likely to choose vocational schools over liberal arts colleges and business or sciences over humanities. As a consequence of these decisions, they frequently benefit from a larger college premium.
Brand and Xie (2010) test these hypotheses by comparing college premiums across propensity score strata. Using data from the National Longitudinal Survey of Youth and the Wisconsin Longitudinal Study, they are able to produce propensity scores that use individuals’ background information to estimate the probability of each individual attending college, regardless of whether or not they actually did attend. Then, comparing individuals with similar probabilities—a process known as propensity score matching—they can see how the magnitude of the college premium varies with the probability of attending college. The results of their analysis provide strong support for the negative selection hypothesis.

Studies exploring differences in the college premium across gender and race provide additional evidence for the debate between positive and negative selection. One such study, conducted by Diprete and Buchmann (2006), uses data from the Current Population Survey to analyze trends in the college premium separated by gender. Using data from 1964 to 2002, they consistently find higher college premiums for women than for men.
Looking at more recent data, Hubbard (2011) finds no statistically significant difference in college premiums between genders, arguing that the difference has been insignificant since around 2000. Given that men were more likely to possess a college degree throughout the period analyzed by Diprete and Buchmann, but not, at least among younger cohorts, more likely to possess a degree during the years studied by Hubbard, these findings would seem to support the negative selection hypothesis. However, it is important to note that female students have historically outperformed their male counterparts by most academic measures, which could provide an alternate explanation for the gender disparity identified by Diprete and Buchmann (Holmlund and Sund, 2008).

Arcidiacono, Bayer, and Hizmo (2010) provide additional insight into the issues at hand with their work on racial discrimination. They find strong evidence of the role of ability signaling for college graduates, noting that such signaling does not take place as efficiently for high school graduates. Consequently, little to no race-based wage discrimination was
found among college graduates in their sample, whereas significant discrimination was apparent in the high school cohort. The authors interpreted these results as showing how the ability signaling effect of higher education largely prevents employers from differentiating wages of college-educated employees based on race, whereas employers looking to hire people with only a high school education have less objective criteria on which to base wages and thus tend to discriminate. A corollary to their findings is that the college premium is higher for blacks than whites. Because college enrollment rates are lower for blacks, this provides support for the negative selection hypothesis.

All three of the studies presented in this section find evidence of the negative selection hypothesis. Yet, it is important to note that the studies by Diprete and Buchmann (2006) and Arcidiacono, Bayer, and Gizmo (2010) only focus on one demographic characteristic each—gender in the case of the former and race the latter—which makes them imperfect tests of negative selection. The original study by Brand and Xie (2010) is superior to the other two due to their use of propensity score matching to look
at numerous demographic characteristics simultaneously. However, propensity score analysis is by no means a perfect method for determining causality, as it suffers from numerous limitations and certain ambiguities in its implementation (see Bai, 2011 for a discussion of these). Accordingly, any conclusion based exclusively on this technique must be taken with a grain of salt.

In sum, the existing literature provides modest evidence for the validity of the negative selection hypothesis. This study finds no evidence whatsoever in support of the positive selection hypothesis. One of the objectives of this thesis is to determine whether the negative selection hypothesis is consistent with trends in college premiums over the 1971-2014 period.

**Age, College Premiums, and the Baby Boomer Generation**

The question of how college premiums vary with age has been addressed in several papers, although there is some disagreement as to the nature of this relationship. Blackburn, Bloom, and Freeman (1991) find no correlation between age and
college premiums in 1979 data, but do find that the college
premium was higher among younger cohorts in 1988. In contrast,
Card and Lemieux (2001) analyze data from 1959 to 1996 that
suggest college premiums were higher for older cohorts during the
majority of the period under study. Looking at the data in selected
three-year chunks, they find that in 1979-1981 and 1984-1986 the
magnitude of the college premium increased with age. In
1989-1991, the relationship became less clear, with college
premiums rising only very slightly with age. Then, in 1994-1996,
the relationship appears to be polynomial: 36-40 year olds
exhibited the largest college premiums, with cohorts on either side
showing smaller premiums. The authors note that over the
1959-1996 period, premiums increased much more for younger
cohorts than they did for older cohorts. This finding is the one
point of commonality between their study and the findings of
Blackburn, Bloom, and Freeman.

In his study on income inequality, Levy (1992) presents an
argument implying that the movement of the baby boomer
generation through age cohorts might have been an important
determinant of the relationship between college premiums and age. In addition to being larger than previous generations, baby boomers also had higher levels of educational attainment (Digest of Education Statistics, 2012). This led to a steep increase in the supply of college-educated workers, which would have disproportionally decreased the college premium among baby boomers.¹ As a result, one would expect the age cohort containing the baby boomer generation to have lower-than-normal college premiums. A subsequent section of this thesis will investigate the relationship between age, college premiums, and the potential impact of the baby boomer generation.

Trends and Determinants of the College Premium

Debating Trends in the College Premium

Freeman and Hollomon (1975) were some of the first economists to start mapping the trajectory of college premiums over time. They concluded that college premiums, after rising steadily in the 1950s and 1960s, experienced a precipitous drop in

¹ This assumes there is not perfect substitution between workers of different ages.
the 1970s. This led some economists to believe that the value of higher education had reached a state of persistent decline, and that trends in college enrollment would follow (Witmer 1976).

These conclusions, especially those of Freeman and Hollomon, sparked a fierce debate regarding the trend in college premiums and the future economic value of higher education. Witmer (1976) argued that premiums had not significantly decreased as of 1975, but were likely to do so dramatically between 1976 and 2000. In contrast, Ostar (1975) found it unlikely that the value of higher education would continue to fall, attributing Freeman and Hollomon’s discovery to poor macroeconomic growth in the first half of the 1970s. Ostar went on to criticize them for ignoring outcomes such as the steadiness of employment, vacation days, and benefits when assessing the value of college. Additional criticisms of Freeman’s (1975) work were put forth by Schwartz and Thornton (1980), who argued that he had inadvertently shrunk his college premium estimates by making a
methodological error in the way they discounted income streams over time.\textsuperscript{2}

Complicating the debate still further, Wish and Hamilton (1980) disagreed with Schwartz and Thornton and argued strongly in favor of Freeman’s models, even recommending their use by university administrators and policymakers in the future. In addition, Freeman (1980) unsurprisingly concurred with Wish and Hamilton in a rebuttal to Schwartz and Thornton. Even today, there seems to be a lack of clarity regarding the trend in college premiums between the late-1960s and the late-1970s.

The economic value of college education continued to be an object of debate in the 1980s and 1990s. The passage of time ultimately proved Ostar (1975) right and Witmer (1976) wrong, as subsequent studies found that college premiums climbed rapidly after 1980. Murphy and Welch (1992) found that the college premium for male workers across all fields and experience levels

\textsuperscript{2} According to Schwartz and Thornton, Freeman’s (1975) college premium estimates are biased by the fact that his income data are in constant dollars, even though he applies a nominal discount rate to income streams over time. The result is that the income earned by college graduates towards the end of their careers is essentially discounted twice.
increased from 38% in the late-1970s to 58% in the late-1980s.

Similarly, Goldin and Katz (2007a) estimate that college premiums rose from around 40% in 1980 to around 55% in 1990, rising further still to 60% in 2000. Studies that consider college premiums over a longer period, such as those of Goldin and Katz (2007a) and Cunha, Karahan, and Soares (2011), have mostly reinforced the finding that premiums were falling in the first half of the 1970s, while discrediting Freeman and Hollomon (1975) and Witmer’s (1976) predictions of persistent declines in the college premium.

Goldin and Katz’s paper is noteworthy because it analyzes the longest period of any study that I have encountered, examining data from 1915 to 2005. Their estimates indicate that college premiums increased from 1980 to 2005, although the rate of this increase slowed over the 25-year period. This is consistent with the findings of Cunha, Karahan, and Soares (2011), even though the two studies use complete different data.

In my opinion, Goldin and Katz present the most compelling study on college premiums over time. One consequence of studying such a long period, however, is that data limitations from
the first two thirds of the 20\textsuperscript{th} century prevent Goldin and Katz from modeling wages as precisely as they could have if they had restricted their sample to more recent years. Several key determinants of earnings are lacking from their model, including race (apart from white versus non-white), marital status, population density, Hispanic origin, and parents’ socioeconomic status. A principal objective of this thesis will be to build on the work of Goldin and Katz and remedy these shortcomings by utilizing a more complete model.

\textit{Identifying Determinants of the College Premium}

In addition to identifying trends in college premiums, studies have looked at what forces may be driving these changes. Supply and demand frameworks frequently form the econometric foundation of such studies. One must turn again to Freeman (1975) for the first example. Analyzing the college job market, he found that fluctuations in the demand for college-educated labor, occurring as supply was steadily increasing, are primarily
responsible for the trends in college premiums in the 1950s through the first half of the 1970s.

More recently, Goldin and Katz (2007a) have used a similar supply and demand model to look at college premiums. Interestingly, their models indicate that fluctuations in supply, not in demand, are the major forces underlying the changes in the college premium. There are a number of reasons why, given their similar methodological approach, they arrive at a different conclusion than Freeman. Goldin and Katz look at a much longer period—1915 to 2005—which allows them to focus on secular shifts in the supply and demand of college-educated labor. Additionally, they factor institutional changes into their model, including variables that account for the inflation and oil price shocks of the mid- and late-1970s (which occurred simultaneously with a marked reduction in productivity growth). These extraordinary circumstances, coming at the end of the period studied by Freeman, may have biased his findings. For that reason, it seems that Goldin and Katz’s analysis is superior to that of Freeman’s.
Looking at supply and demand from a different angle, Acemoglu (1998) investigates the link between the supply of highly educated workers and the demand for what he calls “skills-complementary technology”. Acemoglu argues that an increase in the supply of highly educated labor has two consequences for the wages of highly educated workers: a short-term decrease and a long-term increase. The short-term decrease stems from the textbook instance of excess supply of labor. However, this increase in supply also creates additional pressure to make these highly educated workers more productive by developing technologies that complement their skills. As these skills-complementing technologies are invented, highly educated workers become more productive, leading to an eventual increase in their wages.

Acemoglu’s theory is consistent with the trends identified by Goldin and Katz (2007a) and Cunha, Karahan, and Soares (2011). There was a spike in the number of college graduates in the 1970s, which was accompanied initially by a short-term fall in the college premium and then by a steadily increasing premium in subsequent decades.
Other authors have tested completely different hypotheses with respect to the trends in college premiums. Rumberger (1984) found evidence of a positive correlation between college premiums and macroeconomic growth. It is certainly true that, as college premiums were supposedly shrinking between the late-1960s and the mid-1970s, the United States was going through a period of poor economic growth. It is equally true that prior to this period, when college premiums are thought to have been growing, macroeconomic conditions were healthier. However, the fact that Rumberger only analyzed 20 years of data (1960-1980) made it somewhat difficult for him to defend this theorized correlation.

Some studies have focused on the magnitude of the selection bias, testing to see if the increase in college premiums since the early-1980s can be explained by a growing difference in natural ability between high school and college graduates. Cunha, Karahan, and Soares (2011), for example, analyze the economic returns to certain cognitive skills that are highly concentrated among college graduates. Their findings indicate that these returns...
did increase from the 1970s to the 1980s, which could partially explain the dramatic increase in college premiums they found during this period. However, their model suggests that there was essentially no selection bias in the 1970s, a claim that I find to be highly suspect.

Grogger (1995) also examines whether the selection bias played a role in the increase in college premiums between the 1970s and 1980s. Using data on standardized test scores and high school grades of college graduates, he finds no significant changes among men regarding skills attained prior to college. However, there was a significant increase in math skills obtained during high school among female graduates, which may have increased the magnitude of the selection bias and contributed to the rise in college premiums identified in the studies mentioned earlier (Goldin and Katz, 2007a; Cunha, Karahan, and Soares, 2011).

Also relevant to this discussion is the popularity of different undergraduate majors. There is little doubt that what one studies in college has a significant impact on future income (Oreopoulos and Petronijevic, 2013; Webber, 2014). Graduates with degrees in
business and STEM fields—science, technology, engineering, and math—can expect to earn 1.4-1.5 million dollars more over the course of their lives than ceteris paribus individuals with high school degrees, whereas as the bonus for humanities graduates is about $700,000 (Webber, 2014). Thus, it follows that a shift in enrollment from low-earning degrees to high-earning degrees, or vice versa, could influence the magnitude of college premiums. In fact, Grogger (1995) finds that among men, roughly 25% of the increase in college premiums in the 1980s can be explained by the increased popularity of degrees associated with higher lifetime earnings.\(^3\)

Institutional quality is another factor worthy of consideration. Previous research indicates that college premiums may be associated with the quality of educational institutions (Pascarella and Terenzini, 2005). Institutional quality can be measured in several ways, including dollars spent by the institution per student,\(^3\)

\(^3\) Making this relationship more complex, however, is the finding that the relative popularity of different majors is partially determined by labor market conditions. In a difficult labor market, students tend to pursue courses of study that lead to better economic prospects, whereas in times of low unemployment, less career-oriented majors become more popular (Paulsen and Pogue, 1988).
independent ratings, acceptance rate, the faculty-student ratio, or some combination of these. If there were long-term trends in the institutional quality of colleges in the U.S., this could have an impact on college premiums. More specifically, if metrics such as dollars spent per student or the faculty-student ratio were driving changes in the college premium, one could interpret that as a change in the level of productivity enhancement (see the beginning of this chapter for an explanation of productivity enhancement). In contrast, if the college premium was responding to changes in acceptance rates—a measure of selectivity—then one would be more likely to conclude that the shift in the college premium was coming from the magnitude of the selection bias.

Yet another potential determinant of college premiums is immigration. Since recent immigrants possess less education on average, an increase in immigration rates would raise the relative supply of less-educated labor, thereby producing downward pressure on wages for less-educated workers and increasing the college premium. However, some studies have shown that this downward pressure is offset by the fact that less-educated native
laborers tend to specialize in different occupations than immigrant laborers, even though both hold jobs that don’t require extensive education (Peri and Sparber, 2009). This has the effect of reducing competition and neutralizing the predicted decrease in wages.

Trends in the average amount of hours worked are also worth taking into account when considering changes in the college premium. One study has suggested that a significant portion of college premiums can be explained by the longer hours kept by college graduates (Zhang, 2008). If the length of an average work week were to change for either high school or college graduates, one would expect such a change to have an impact on college premiums.4

Unionization may also have an effect on college premiums. This hypothesis is based on the fact that union jobs are more likely to be held by workers without a college degree. In addition, one of the primary objectives of unions is to raise real wages for members, many of whom work for the minimum wage. If union

---

4 Zhang’s (2008) study also found that the amount of hours worked varies by undergraduate major and by industry, thus connecting it to Oreopoulos and Petronijevic’s (2013) aforementioned work.
membership were to fall, one would expect there to be downward pressure on real wages. Since this downward pressure would be concentrated among workers without a college degree, college premiums would increase. This is indeed what may have happened over the last half-century: between 1973 and 2007, union membership dropped from 34% to 8% among men and from 16% to 6% among women, and the real minimum wage experienced a significant decline (Western and Rosenfeld, 2011). This is consistent with the studies indicating that college premiums have risen since the 1980s (Goldin and Katz, 2007a; Cunha, Karahan, and Soares, 2011). Adding to this, Lindley and Machin (2013) find that states with higher union membership tend to have lower college premiums. Taken together, these studies present compelling evidence for the link between college premiums and unionization.

A recent study by Guvenen and Ozkan (2014) provides yet another perspective on the trend in college premiums by examining the role of tax policy. Their argument begins with the widely accepted premise that progressive taxation discourages
investment in human capital by reducing the anticipated post-tax income gains that typically motivate human capital investments. What is novel about their work is the finding that high-ability individuals appear to be more discouraged from investing in human capital than low-ability individuals. Thus, progressive taxation reduces the selection bias, which in turn lowers the estimates of college premiums. The fact that the U.S. has maintained a relatively non-progressive tax policy since the 1980s is consistent with the increase in college premiums found by other authors (Goldin and Katz, 2007a; Cunha, Karahan, and Soares, 2011).

One final point to consider when discussing the determinants of college premiums is the Heckscher-Ohlin Model. A foundational set of theories in the study of international trade, the Heckscher-Ohlin Model predicts that a country open to trade will export the goods and services that intensively require the country’s relatively abundant factors of production. The Stolper-Samuelson Theorem, a major component of the Heckscher-Ohlin Model, adds to that prediction by arguing that the owners of the abundant
factors will benefit, as demand will rise for the factors they possess. In the case of the United States, the Heckscher-Ohlin Model predicts that the U.S. will mostly export goods and services that require high levels of human capital—in other words, goods and services that require highly educated labor. Thus, according to the Stolper-Samuelson Theorem, one would expect an increase in international trade to benefit the owners of human capital (the highly educated workers themselves), which would lead to higher college premiums. It should be noted, however, that the Heckscher-Ohlin Model is often not substantiated by looking at real trade relationships (Deardorff, 1982). More recent applications of Heckscher-Ohlin have added depth to the model by allowing for more complex wage structures and differing ability levels, but the connection between the Stolper-Samuelson Theorem and college premiums remains uncertain (Haskel et al., 2012).

There are many potential explanations for the trends in college premiums. Some of these explanations are mutually exclusive, but many are just as valid in tandem with others as they are when viewed alone. A primary objective of this thesis will be to
test the explanatory power of some of the theories discussed in this chapter.
Chapter 2: Trends in College

Premiums, 1971-2014

The primary goal of Chapter 2 is to answer the following questions. First, what is the current college premium in the United States, and how has it changed over time? Second, to what extent do college premiums vary across gender, race, and ethnicity? Although previous studies have investigated these issues separately, Chapter 2 offers a comprehensive and integrated analysis of college premiums over the last 44 years. It also introduces several innovations aimed at increasing the accuracy and precision of college premium estimates.

Data

To investigate these questions, data on individuals from the Current Population Survey (CPS) are used. The basic CPS is conducted on a monthly basis, but every March an additional
segment called the Annual Social and Economic Supplement (ASEC) is included. The ASEC is the primary data source for this study. Although the CPS was established in 1940, it has gone through a number of structural and methodological iterations, making comparisons between years problematic. However, with detailed data management, it is possible to modify data from 1971 to 2014 so that they are comparable over time.

It should be acknowledged that some issues persist even after data management. One problem with comparing CPS data over time is that the sample weights of different demographic groups have been intentionally varied at numerous points over the last 50 years (King et al., 2010). Although running weighted regressions allows one to sidestep any potential biases created by this practice, inconsistent sampling still has an effect on the size of

---

5 The most frequent difficulty in comparing CPS data over time is that answer choice sets are subject to change between years. For example, between 1971 and 2014, the response options for the question assessing race were modified five times, growing from 3 options in 1971 to 25 options in 2014. Another common issue stems from the way questions are framed. For example, prior to 1992, educational attainment was assessed in terms of number of years of completed schooling. After 1992, the education question was replaced with one asking respondents to indicate their highest level of competed education (i.e. high school, college, etc.).
standard errors. This is not a major issue in my estimation, but it should still be noted as a drawback to the data used here.

Another weakness of CPS data is that they are “top coded”, so that respondents who report incomes over a certain amount are removed from the publicly accessible data. The logic behind this technique is to ensure anonymity, given that respondents with sufficiently high earnings may be identified by the size of their incomes. A downside of top coding, as noted by Hubbard (2011), is that it not only compresses estimates of college premiums, but it may do so to a greater degree for men than for women. This is because a disproportionately high number of male respondents are top coded, the majority of whom possess college degrees. Although top coded CPS datasets are frequently used in academic research, this shortcoming should be kept in mind when analyzing the results in this chapter.

For the purpose of this thesis, the datasets were trimmed so as to only include survey respondents whose highest level of completed education was either college or high school. Respondents with partially completed levels of education—i.e.
“some college”—were dropped from the sample as well. The sample was modified to consist exclusively of respondents in the labor force, either as self-employed workers or as employees. Respondents who were unemployed, not in the labor force, working as an unpaid family worker, or in the armed services were not included in the sample.

The economic outcome used to estimate college premiums is annual income. Two forms of income are included: income earned from wages and salary, and income earned from non-farm self-employment. All other income sources—self-employed farm income, capital gains, social security, alimony payments, et cetera—were excluded from this analysis.

**Methodology**

The first step in estimating college premiums was to produce models for annual income using variables available in the CPS and ASEC. Included in this model were variables for gender, age (included as a level and a quadratic term), marital status,
Hispanic, Latino, or Spanish origin, region of the U.S., and population density of the respondent’s place of residence.

Accounting for race presented something of a challenge, as there were only three racial categories in the 1971 CPS (white, black, and “other”). To maintain comparability with later data, the same three categories are used throughout the analysis, with all respondents not identifying as only white or only black categorized as “other”.

The educational attainment of each of the respondents’ parents was also included as a proxy—admittedly imperfect—for socioeconomic background. Lastly, a dummy variable separating the sample between high school and college graduates was added. Annual income was used as the outcome variable, and population-weighted multivariate regressions were run for each year of the sample.\(^6\)

As stated previously, the definition of the college premium used throughout this thesis is the mean percentage difference in income between college and high school graduates. Estimates of the college premium in a given year are taken from the coefficient

\(^6\) Annual income was used in logged form throughout this analysis so as to identify changes in percentage form.
on the education variable. An estimated coefficient of 0.4, for example, would indicate that possessing a college degree is associated with a 40% increase in annual income. In order to look at differences in college premiums between groups—for example, men versus women—an interacted term was added to the model consisting of the education variable and the group identifier (in the case of gender, a binary variable for sex). All of the estimates discussed in this chapter are displayed in the Appendix, found at the end of this thesis.

**Results and Discussion**

**Full Sample Estimates**

Graph 1 presents the estimated college premiums for each year of the sample. As expected, the education variable is highly significant in every year. The college premium appears to decrease over the first five years of the sample, starting at 35% in 1971 and hitting a minimum of 26% in 1975. After a rebound in 1976, the college premium fell slightly for the remainder of the decade. It increased rapidly during the 1980s, reaching 47% in 1990.
Note: Black dots represent college premium estimates, thin dotted lines represent 95% confidence intervals.
Continuing to grow during the 1990s, the college premium arrived at a value of 50% in 2000. It remained constant at around 50% during the first 7 years of the new millennium, after which point it began to increase again, reaching a maximum of 57% in 2012.

Confidence intervals set to 95% are plotted around the estimates in Graph 1. They suggest that the model is able to estimate college premiums with a relatively high degree of precision, especially after the turn of the century. Although the confidence intervals are narrowing over the 44-year period, this can likely be attributed to increasing sample sizes rather than bias in the model.

To check whether or not data was successfully managed so as to be comparable over time, one can look at the volatility of estimates between years in which changes were made to the CPS, and compare that with the volatility between years in which no changes were made. Such an examination yields generally encouraging results, with transition years actually exhibiting less volatility than non-transition
years.\(^7\) There is one notable exception, which is 1976. This is the first year in which Hispanic respondents were oversampled in the CPS (King et al., 2010). There was also a change in how Hispanic, Latino, and Spanish origins were assessed in the survey. Although the models used in this study do control for Hispanic, Latino, and Spanish origin, these changes could have caused a bias in the results if they led people to identify as Hispanic, Latino, or Spanish in 1976 who would not have done so in 1975 (or vice versa). Although it is important to acknowledge the possibility of this bias, there is no direct evidence indicating it exists, and it is also possible that the spike in college premiums in 1976 is simply coincidental. As is discussed later on, this latter interpretation appears more probable when separating college premiums by gender and ethnicity.

Methodological differences make it hard to compare the results presented here with those of similarly focused studies, but the estimates shown in Graph 1 are generally consistent with pre-existing

\(^7\) To look at average volatility between transition and non-transition years, the absolute values of the differences between years were calculated and averaged for the two groups. The results were 1.601 for transition years (1975-1976, 1987-1988, 1991-1992, 1995-1996, 2002-2003, 2012-2013, and 2013-2014) and 1.97 for non-transition years, indicating that volatility was actually lower in transition years.
work aimed at mapping college premiums over time. The study by Goldin and Katz (2007a) is perhaps most similar in terms of execution to the one presented here. The trends identified in their paper—decreasing college premiums in the 1970s, rapidly increasing premiums in the 1980s, followed by less rapid increases from 1980-2005—closely match the trends identified here. This pattern is also found in the work by Cunha, Karahan, and Soares (2011).

In terms of the magnitude of college premiums, Goldin and Katz’s results are moderately higher than the ones presented here throughout the 35 years of overlap between the two studies. On the other hand, the results of this thesis match up closely with the estimates put forth by Cunha, Karahan, and Soares. There are several potential explanations for the difference between the estimates presented here and those of Goldin and Katz. First, their analyses only include employees, whereas this thesis includes employees and non-farm self-employed workers. It is possible that the self-employed workers in the sample used here put downward pressure on college premiums, causing the disparity noted above. In fact, back-of-the-envelope calculations on my data show that the college premium for
employed workers is higher than the premium for self-employed workers in 37 of 44 years. Another potential explanation is that the model employed in this thesis controls for more confounding variables than the model used by Goldin and Katz, and many of the variables missing from their model tend to bias upwards the college premium. For example, removing variables for marital status, Hispanic origin, and population density from my models and rerunning the regressions from 1990 to 2005—the last 15 years of their sample period—leads to an increase in the estimated college premium of two to four percentage points. These variables were not included in the paper by Goldin and Katz, likely due to the limitations explained in the previous chapter. Thus, it seems probable that differing model specifications are partially responsible for the differences in college premium estimates. It is my

---

8 Data were separated between employees and self-employed workers, and regressions identical to the ones used throughout this study were run on each group. The college premium estimates for the self-employed group had very large standard errors because the sample sizes were comparatively small. For this reason, I decided to use OLS regressions instead of population-weighted regressions. Although this is not ideal, the fact that the premium estimates for the self-employed cohort were lower than the employee cohort in 37 of 44 years does suggest that college premiums are smaller for self-employed workers.
belief that the models used here are more accurate, seeing as they account for a larger number of key determinants of income.

**Gender Differences in College Premiums**

Graph 2 uses an interaction term to separate college premiums by gender. The red squares indicate years in which the difference between the two estimates was statistically significant. As is evident, there is considerable divergence between genders prior to the turn of the century, yet this divergence is not uniform. For the first half of the 1970s, women exhibit a drastically higher college premium than men. However, starting in 1976 and continuing until 1984 this difference is statistically insignificant, as college premiums begin to rapidly rise for both genders. The difference is once again significant for the majority of the 1985-1997 period, with college premiums rising much more quickly for women between 1984 and 1994. After this point, there is a decline in women’s college premiums that causes the gender difference to become insignificant starting in 1998, and it remains so until the end of the period. Interestingly, over the final 20 years of the period (1995-2014), men’s college premiums rose by 13 percentage
Note: Red squares indicate a statistically significant difference (95% CI) between men and women.

Graph 2: College Premiums, Men and Women (1971-2014)
points while college premiums for women did not increase at all.

These findings are fairly consistent with the conclusion put forth by Diprete and Buchmann (2006) that college premiums were higher for women between 1964 and 2002. The biggest contrast in our findings occurs between 1976 and 1984, when my results indicate no significant difference between genders. This dissimilarity may be due to the fact that they used a very different sample than I did. Their data were confined to white, full-time workers aged 30-34, whereas I excluded no respondents on the basis of race or age. My findings are also consistent with Hubbard’s (2011) conclusion that there has been no gender gap in college premiums since around 2000. The results presented here can be seen as an expansion of Hubbard’s analysis, given that he used a much more homogeneous sample (similar to Diprete and Buchmann): white, non-Hispanic employees with full-time, full-year jobs that earned above a certain level of income. Since no such restrictions were made to the data analyzed in this thesis, the finding that the gender gap in college premiums has been nonexistent since the turn of the century can be extended to a much larger portion of the U.S. population.
One salient difference between my estimates for men and women is that college premiums for men follow a much more linear path over the 1971-2014 period. To look at this mathematically, lines of best fit were produced for both genders. The $R^2$ value of the fitted line for men (0.94) was significantly higher than for women (0.64), illustrating this difference. Unlike the trend in men’s college premiums, the trajectory of college premiums for women has two clear troughs, one around 1980 and the other around 2004. It also has a definite peak around 1994.

It should be noted that the spike in college premiums in 1976, which was discussed earlier as potential evidence for a lack of comparability between 1971-1975 data and 1976-2014 data, seems to be almost entirely concentrated among men. This observation is inconsistent with the explanation related to sampling and coding procedures for the Hispanic population. Since there were no methodological changes to the CPS in 1976 that would have disproportionately affected male respondents, I find it likely that the spike in college premiums is purely coincidental.
Race and Ethnicity Differences in College Premiums

Graph 3 separates college premiums between whites and blacks. Due to the limitations mentioned earlier regarding how the CPS assessed race in the 1970s, white and black are the only two racial categories that can be modeled individually over the entire 1971-2014 period. Consequently, respondents who identified with other races—including mixed race respondents who included white or black in their racial identification—are dropped from the sample when estimating the college premiums shown in Graph 3. Blacks exhibit higher college premiums in all 44 of the included years, with the difference reaching statistical significance in 35 of these years. This finding is consistent with the results of Arcidiacono, Bayer, and Hizmo (2010). The magnitude of the difference appears to have decreased slightly over time, averaging 16 percentage points in the first half of the period and 12 percentage points in the second half. Compared to the male-female difference, however, the divergence presented in Graph 3 is notably consistent over time. The increased volatility in the estimates for blacks stems from a relatively small black subsample, and is not likely a sign of bias.
Note: Red squares indicate a statistically significant difference (95% CI) between blacks and whites.
In Graph 4, individuals of Hispanic, Latino, or Spanish descent are separated from the rest of the sample. The graph suggests that there is little to no difference in college premiums between the two groups. Apart from what appear to be outliers in 1972 and 1977 and a two-year stretch from 2002-2003 in which the difference was statistically significant, the estimates for the Hispanic, Latino, and Spanish population seem to fluctuate closely around the estimates for the rest of the sample. As with the previous graph, the increased volatility in the Hispanic, Latino, and Spanish estimates can be explained by a smaller subsample size.

The reason for the statistically significant difference in 2002-2003 remains unclear. There were no methodological changes made to the CPS in either 2002 or 2004 that seem likely to have influenced college premium estimates for the Hispanic, Latino, or Spanish population (King et al., 2010). Although it is possible that the immigration spike in the late-1990s is responsible, by 2002 net migration rates had already been decreasing for a few years (World Bank, 2015). Additional investigation of these two atypical years goes beyond the scope of this thesis.
Note: Red squares indicate a statistically significant difference (95% CI) between the two groups.

Graph 4: College premiums, Hispanic, Latino, or Spanish Origin (1971-2014)
The estimates for the Hispanic, Latino, and Spanish cohort provide additional insight into whether or not the spike in 1976 in the full sample estimates is due to methodological changes in the CPS. Since these changes were all concerned with how to sample and code people of Hispanic, Latino, or Spanish descent, one would expect Graph 4 to have some sort of outlier in 1976 if these methodological changes were indeed the underlying cause. No outlier is observed, making it even more probable that the 1976 spike was purely random.

The College Premium Between Age Cohorts

As discussed in the literature review, there exists some disagreement regarding the relationship between age and the magnitude of college premiums. Graph 5 separates college premium estimates into two age cohorts, with age 35 as the cutoff. During the first decade of the period, the older cohort exhibited higher college premiums in every year, with the difference reaching statistical significance in five of ten years. The difference was relatively modest, though, averaging only six percentage points between 1971 and 1980. In contrast, between 1982 and 2014 college premiums were
Note: Red squares indicate a statistically significant difference (95% CI) between the two groups.

Graph 5: College Premiums by Age Cohort (1971-2014)
larger in the younger cohort, and the difference was statistically significant in every year. The magnitude of this difference stayed relatively consistent over the 33-year span, averaging 10 percentage points between 1982 and 1999 and 9 percentage points between 2000 and 2014.

These results are similar to those of Blackburn, Bloom, and Freeman (1991), but are not entirely consistent with the findings presented by Card and Lemieux (2001). Analyzing data from 1959 to 1995, the latter authors find a positive correlation between age and college premiums in the 1970s and 1980s. This changed in the middle of the 1990s, when their results begin to show 36-40 year-olds exhibiting larger college premiums than both younger and older cohorts. There are numerous potential explanations for why Card and Lemieux’s findings contrast with mine. First, they neglect to control for the polynomial nature of the relationship between age and earnings. This is accomplished in my models by controlling for age with both level and quadratic terms, in addition to interacting the age cohort variable with the education binary. The fact that the coefficients for both the level and the quadratic terms in my models are statistically
significant in every year demonstrates the importance of modeling age in this way. Not doing so likely resulted in the coefficients on Card and Lemieux’s age-education interacted terms capturing some of the relationship between age and income, even though this relationship should have been fully separated out to avoid introducing a bias.

Additionally, Card and Lemieux employ a much simpler model than the one used here. In my estimation, they leave out some crucial determinants of income—marital status, Hispanic origin, and regional variables, to name some examples—which could have subjected their estimates to numerous biases. For these reasons, and because my findings are consistent with those of Blackburn, Bloom, and Freeman, I believe the results presented here are more accurate than those of Card and Lemieux.

The estimates shown in Graph 5 are consistent with Levy’s (1992) study in that they suggest the baby boomer generation may have had a substantial effect on the relationship between college premiums and age. As noted in the literature review, the baby boomers were different than previous generations in two relevant ways: first, their generation was much larger than previous
generations, and second, they were more likely to pursue a college education than their predecessors. This led to a spike in the supply of college-educated workers in their age range, which likely put downward pressure on the college premium for their generation.$^9$

The birth years of baby boomers are concentrated between the late-1940s and the early-1960s. This means that they would be part of the young cohort until the early-1980s, at which point they would start to turn 36 and join the older cohort. The transition would be complete sometime in the mid-1990s, when almost all of the baby boomers had turned 36. Considering baby boomers’ hypothesized impact on college premiums, one would expect the downward pressure to fall entirely on the younger cohort prior to the early-1980s and entirely on the older cohort after the mid-1990s, with an ambiguous effect in the intermediate years. This is consistent with Graph 5, in which college premiums are lower for the younger cohort in the 1970s and lower for the older cohort from the 1980s on.

---

$^9$ This statement assumes that workers of different ages are not perfect substitutes.
Chapter 3: College Premiums and the Economic Value of Higher Education

The results presented in Chapter 2 indicate that college premiums have increased over the last 39 years. This does not necessarily mean, however, that the “economic value” of going to college has increased. In this thesis, college premiums are defined as the mean percentage difference in income between college and high school graduates. In contrast, the economic value of higher education can be thought of as the increase in income that stems directly from possessing an undergraduate degree.

I find two principal differences between the college premium and the economic value of an undergraduate degree. The first, outlined in the literature review, is selection bias. In order to be accepted into and graduate from college, one must possess a certain level of “natural” or pre-college ability. Graduation from college is
therefore not random among high school graduates: on average, the more accomplished and competent high school graduates end up selecting into higher education. As a consequence, it is likely that individuals who possess an undergraduate degree would be more valuable in the workplace than those who do not, even if the college graduates had never gone to college. This is because the same skills and abilities that cause individuals to select into college also make them more valuable in professional settings. Because they are more valuable, they tend to earn higher wages, which is how the selection bias pushes up college premiums without actually adding to the economic value of going to college.

The second difference relates to the popularity of different college majors. There is substantial heterogeneity in the future earnings associated with different majors, with business and STEM programs—science, technology, engineering, and math—leading to much higher lifetime earnings than humanities (Oreopoulos and Petronijevic, 2013; Webber, 2014). Therefore, a shift in relative enrollment rates between low-earning majors and high-earning majors should have an effect on college premiums. However, it is my opinion
that this would not reflect a change in the economic value of higher
education, so much as a change in the kind of higher education that
was being pursued.

This chapter aims to bridge the gap between college premiums
and the economic value of higher education by assessing the dual
impacts of selection bias and major choices. I will focus on the
following question: given that college premiums rose over the
1975-2014 period, can we conclude that the economic value of going
to college increased as well?

The Magnitude of the Selection Bias Over Time

Without time-consistent data on the pre-college ability of
individuals, it is impossible to measure and control for the selection
bias. It is possible, though, to infer how the magnitude of the
selection bias has likely changed over time. When considering how
this bias shifted over the last 40 years, it is important to keep in mind
several facts. First, both the percentage of the population with a
bachelor’s degree and the percentage of the population with a
graduate degree steadily increased between 1975 and 2014 (Digest of Education Statistics, 2012). Even so, the percentage of the population with a bachelor’s degree or higher remained well under 50%, reaching only 32% in 2014 (U.S. Census Bureau, 2014).

Image 1 illustrates how these trends have affected the selection bias. It is assumed that “natural ability”, as discussed in the opening section of this chapter, is normally distributed in the population. It is also assumed that educational attainment and natural ability share a strong positive correlation. In the beginning of the 1975-2014 period, when the percentage of the U.S. population possessing postsecondary education was low compared to later on, it is arbitrarily
decided that areas C and D represent the portions of the population whose highest level of completed education was an undergraduate degree and a graduate degree, respectively. Areas A and B, therefore, represent the portion of the population possessing only a high school degree. Accordingly, the selection bias in college premiums stems from the difference in average ability between the people in area C and the people in area A+B.

Moving to the end of the 1975-2014 period, the percentage of the population possessing an undergraduate degree has risen, along with the percentage of the population possessing some form of graduate degree. In Image 1, this translates to areas C and D now representing the portion of the population with a graduate degree and area B representing the portion of the population with an undergraduate degree. This shift relies on the aforementioned assumption that educational attainment and natural ability are strongly correlated. With these changes, the new selection bias comes from the difference in average ability between individuals in area B and individuals in area A. This contrasts with the beginning of the period, where the relevant difference was between C and A+B. Because all of
the changes have taken place above the mean of the distribution, ensuring that the area of A remains larger than the areas of B or C, the difference in average ability between the high school cohort and the undergraduate cohort has become smaller. Consequently, the selection bias will have decreased in magnitude.

This finding is clearer when presented in narrative form. As the percentage of the population with a college degree rises—which is essentially equivalent to a decrease in the selectivity of college as a whole—the college-educated cohort is less and less comprised of the “cream of the crop”. In other words, it becomes easier for students of more average ability to get a college education, thereby pulling down the mean ability of the college cohort. Furthermore, as an increasing percentage of college graduates go on to grad school, more and more of the highest-ability college graduates enroll in graduate programs, removing themselves from the college cohort and lowering still further the college cohort’s mean ability. In sum, the college cohort loses individuals with higher ability while gaining more individuals with average ability, therefore experiencing a reduction in mean ability. Although the mean ability of the high school cohort also
decreases when college enrollment increases, this change should be smaller in magnitude since the high school cohort is much larger than the college cohort to begin with.

As mentioned earlier, this theory relies on several assumptions. Although most of them seem reasonable, the assumption that educational attainment and natural ability are strongly correlated does lead to some complications. Among other things, this implies that colleges are highly capable of singling out and admitting the applicants with the highest ability. It must be noted, however, that doing so with 100% accuracy is clearly impossible. This suggests one major unanswered question regarding the theory presented here and the implication that the selection bias has decreased over the last 40 years. If colleges became better over time at assessing ability and admitting only the most intelligent and capable applicants, this would have put upward pressure on the magnitude of the selection bias. Although this does seem plausible, I am aware of no research that supports such a conclusion. Additionally, for the selection bias to have increased overall, this ability-identifying effect would need to have
been powerful enough to outweigh the selectivity effect described with Image 1.

Another critical assumption in this theory is that the shape of the natural ability distribution has stayed relatively consistent over time. If the kurtosis or skewness of the distribution changed, this would introduce additional complications to my line of reasoning. However, I know of no studies that suggest such changes did take place between 1975 and 2014, and the shape of the ability distribution would have to have changed substantially for my theory to no longer be valid.

To conclude this section, it must be acknowledged that the theoretical argument outlined here would be inferior to a more empirical analysis of the selection bias, if such an analysis were to exist. Although I am not aware of any single source of data that would be adequate for such an analysis, it may be possible to piece together different longitudinal studies from the 20th century to try and determine how the magnitude of the selection bias has changed. The problem with this approach, and the reason it was not attempted in this thesis, is that commonly used measures of ability—SATs, GPA, etc.
—have changed drastically over time. This would render any estimates of the selection bias incomparable between periods. Consequently, this section was confined to a more theoretical analysis, which suggests that the selection bias in college premiums has decreased over the last 40 years.

**Trends in College Major Choices**

This section investigates the extent to which shifts in the popularity of different college majors may have driven changes in college premiums between 1971 and 2014. Grogger’s (1995) paper is the only work I have encountered that explores this relationship. Looking only at men in the 1980s, he finds that 25% of the increase in college premiums can be attributed to the rising popularity of majors associated with higher lifetime earnings. This section will build on his work by analyzing data over a longer period, in addition to looking for differences between age cohorts.

Annual data generated by the National Center for Education Statistics were used. These data were modified to produce time series on the number of degrees granted per year in three different
categories: business, STEM fields, and humanities. These categories are based on Webbers’ (2014) findings that business-related and STEM majors are associated with the highest lifetime earnings and humanities with the lowest. Social sciences, the other big category of undergraduate degree programs, were not included in the analysis since they are average with respect to lifetime earnings and thus not as helpful when considering the relationship between major choices and college premiums. The time series were put into percentage form by dividing the data by the total number of bachelor’s degrees granted in each year.

Graph 6 displays these percentages in each year from 1971 to 2012 (data are not included from 2013 and 2014 due to lack of consistency). The popularity of business-related majors rose between 1971 and 1988, starting at around 14% in the early-1970s and rising to

---

10 The definition of each category is as follows:
- **Business**: Comprised of a single NCES category, titled “Business and Business-Related Fields”.
- **STEM**: Includes the following NCES categories: “Biological and Biomedical Sciences”, “Engineering and Engineering Technologies”, “Computer and Information Sciences”, “Mathematics and Statistics”, “Physical Sciences and Science Technologies”, “Economics”, and “Health Sciences and Related Programs”.
- **Humanities**: Includes the following NCES categories: “English Language, Literature, and Letters”, “Foreign Languages and Literatures”, “Visual or Performing Arts”, “Architecture and Related Services”, and “History”.


Graph 6: Undergraduate Major Choices and College Premiums, 1971-2012

Note: Dashed lines graphed on the right-hand vertical axis, solid lines graphed on the left-hand vertical axis.
almost 25% in the late-1980s. Between 1988 and 1997, major choices shifted away from business, hovering around 20% in the late-1990s. The popularity of business programs rose again starting in 1999, but only slightly, staying between 20% and 22% for the remainder of the period. The popularity of STEM fields also rose during the first half of 1971-2012, starting at 21% in 1971 and rising to 31% in 1985. After this point, enrollment in STEM fields decreased, falling to around 23% in the early-1990s. STEM popularity changed little after this point, fluctuating within two percentage points of 25% until 2012. Unlike business and STEM majors, the popularity of humanities programs decreased during the first half of the period, falling from 20% in 1971 to 11% in the mid-1980s. It increased slightly in the following years, but stayed around 13% for the rest of the period.

College premium estimates are also plotted in Graph 6. The estimates for the entire sample are plotted, in addition to the estimates for the younger cohort (individuals under the age of 36). This is done because the data on college major choices corresponds entirely to recent college graduates. If there is a relationship between major choices and college premiums, it should be more evident by
looking only at the younger cohort, of which recent graduates make up a larger portion.

In this vein, it should be acknowledged that single-year graduation data are not strictly compatible with my college premium estimates, since the estimates are based on samples containing many individuals who had been out of school for some time when they provided their data. Unfortunately, consistent data on undergraduate major choices do not go back far enough in time to estimate the major choice distribution of the entire CPS samples. In fact, by pure coincidence, usable data is only available starting in 1971. In addition, the distinction between bachelor’s and master’s degrees is blurry in several fields prior to the 1970s. For these reasons, single-year data are used, which give an idea of how the popularity of different majors was shifting during the period in which college premiums were estimated.

According to Webber’s (2014) findings, one would expect college premiums to be positively correlated with the percentage of students majoring in business and STEM fields and negatively correlated with the percentage of students majoring in the
humanities.\textsuperscript{11} Graph 6 is mostly inconsistent with these predictions. Between 1971 and the mid-1980s, there is a clear shift away from humanities majors and towards business and STEM programs. This conflicts with the hypothesized relationship, given that premiums were decreasing or stagnant for most of the 1970s. In the first half of the 1980s, however, the trend in college premiums reverses and it does briefly appear as though the shift towards high-earning majors could be having the expected effect on college premiums. After the mid-1980s, though, this ceases to be true. Major choices shift away from business and STEM fields starting in the late-1980s and flatten out between the mid-1990s and 2012. Additionally, the popularity of humanities majors rises slightly between the mid-1980s and 2012. Given the dramatic increase in college premiums between 1980 and 2012, this is clearly inconsistent with the hypothesis.

One could argue that even if the hypothesized relationship were accurate, it might not be evident from Graph 6. One explanation for this could be that there is a time lag between shifts in college

\textsuperscript{11} It is assumed that Webber’s conclusions are applicable to the entire 1971-2012 period, meaning that the relative earnings associated with business, STEM, and humanities majors have stayed fairly consistent over time.
major choices and the corresponding impact on college premiums. A related reason, mentioned previously, is that the data on college majors are only for recent graduates, and thus produce highly imperfect representations of the samples used to generate the premium estimates. Even if these arguments were valid—and they certainly may be—the size of the shifts in major popularity still do not seem nearly large enough to substantially contribute to the increase in college premiums. None of the three categories analyzed here fluctuate outside a range of approximately ten percentage points.

Considering Webber’s (2014) finding that the lifetime earnings bump for business and STEM graduates is no more than twice as large as it is for humanities graduates, it is hard to see how such relatively minor, multi-directional shifts could explain a significant portion of the 40 percentage-point increase found in the younger cohort’s college premiums between 1975 and 2012.\(^{12}\)

---

\(^{12}\) It is important to note that Webber is not saying business and STEM graduates earn twice as much as humanities graduates. Rather, his analysis shows that they benefit from college premiums that are twice as large. This is obviously a much smaller difference than if business and STEM graduates earned twice as much as humanities graduates.
In order to further investigate the relationship between college major choices and college premiums, OLS regressions were run with the college premium as the outcome variable and the three major choice time series from Graph 6 as the explanatory variables. Regression output is summarized in Table 1. Specifications (1) and (3) use data on major choices from individual years, whereas regressions (2) and (4) use moving averages, calculated by averaging a given year’s data with the percentages from the two previous years. This was done in order to minimize the effect of short-term fluctuations. In addition, specifications (1) and (2) use college premium estimates for the full sample as the dependent variable, while specifications (3) and (4) use premium estimates from only the younger cohort.

The hypothesis being tested is that college premiums are positively correlated with the percentage of students graduating with business and STEM degrees and negatively correlated with the percentage of students choosing humanities programs. The coefficient on the business variable—positive and statistically significant in each specification—is the only one that supports this hypothesis. The coefficient on the STEM variable has the opposite-of-expected sign in
each specification and is statistically significant in the specifications in which moving averages were used. The coefficient on the humanities variable also has the contradictory sign in all four specifications, though it is never statistically significant. As is evident, the results from Table 1 fail to support the hypothesis that undergraduate major choices had a significant impact on the trend in the college premiums.

![Table 1: Regression Output, College Major Choices and College Premiums (1971-2012)](image-url)

- **DV: College Premium**

<table>
<thead>
<tr>
<th>Specification:</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>N=42</td>
<td>N=40</td>
<td>N=42</td>
<td>N=40</td>
<td></td>
</tr>
<tr>
<td>Adj. $R^2$=.38</td>
<td>Adj. $R^2$=.48</td>
<td>Adj. $R^2$=.47</td>
<td>Adj. $R^2$=.55</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.747)*</td>
<td>(.966)*</td>
<td>(.949)*</td>
<td>(1.22)*</td>
</tr>
<tr>
<td>STEM Fields</td>
<td>-1.638</td>
<td>-2.321</td>
<td>-1.858</td>
<td>-2.742</td>
</tr>
<tr>
<td></td>
<td>(-0.818)</td>
<td>(.662)*</td>
<td>(-1.039)</td>
<td>(.837)*</td>
</tr>
<tr>
<td>Humanities</td>
<td>1.046</td>
<td>0.513</td>
<td>1.445</td>
<td>0.520</td>
</tr>
<tr>
<td></td>
<td>(-1.539)</td>
<td>(-1.887)</td>
<td>(-1.956)</td>
<td>(-2.383)</td>
</tr>
</tbody>
</table>

Notes on regression specifications:
(1): single year data, full sample
(2): moving averages, full sample
(3): single year data, younger cohort
(4): moving averages, younger cohort
Additional note: standard errors are in parentheses, with "*" signifying statistical significance at the 95% level.
The regressions in Table 1 were also run on the first and second halves of the period individually. The results from 1971-1991 are nearly identical to the results from the full period. The regression output looks much different, however, when only analyzing the second half of the period (1992-2012). Results from these regressions are presented in Table 2. Specifications (1) and (3) show all three coefficients with the expected sign, although only the humanities variable reaches statistical significance. On the other hand, specifications (2) and (4) are highly inconsistent with the hypothesis. The fact that the hypothesis is only supported when single-year data are used—and clearly not supported when moving averages are used—prevents one from interpreting these results as strong evidence for the link between major choices and college premiums. In addition, the majority of the increase in college premiums between 1975 and 2012 took place in the first half of the period, and the first-half regression output closely resembles the results in Table 1. In sum, there is little to no evidence that shifting major choices had a significant impact on the rise in college premiums between 1975 and 2012.
As is apparent, the regressions summarized in Tables 1 and 2 are extremely simple and neglect to control for the numerous determinants of college premiums. In addition, they may suffer from a

<table>
<thead>
<tr>
<th>Specification</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N=21</td>
<td>N=21</td>
<td>N=21</td>
<td>N=21</td>
</tr>
<tr>
<td>Adj. $R^2$</td>
<td>.29</td>
<td>.18</td>
<td>.34</td>
<td>.07</td>
</tr>
<tr>
<td>Business</td>
<td>1.343</td>
<td>-1.842</td>
<td>2.590</td>
<td>-1.464</td>
</tr>
<tr>
<td></td>
<td>(1.076)</td>
<td>(1.269)</td>
<td>(1.403)</td>
<td>(1.811)</td>
</tr>
<tr>
<td>STEM Fields</td>
<td>0.2884</td>
<td>-0.298</td>
<td>1.137</td>
<td>0.176</td>
</tr>
<tr>
<td></td>
<td>(0.600)</td>
<td>(0.597)</td>
<td>(0.782)</td>
<td>(0.852)</td>
</tr>
<tr>
<td>Humanities</td>
<td>-5.278</td>
<td>-2.363</td>
<td>-6.598</td>
<td>-3.426</td>
</tr>
<tr>
<td></td>
<td>(1.660)*</td>
<td>(1.839)</td>
<td>(2.164)*</td>
<td>(2.624)</td>
</tr>
</tbody>
</table>

Notes on regression specifications:
(1): single year data, full sample
(2): moving averages, full sample
(3): single year data, younger cohort
(4): moving averages, younger cohort
Additional note: standard errors are in parentheses, with "*" signifying statistical significance at the 95% level.
simultaneity issue, as noted by Paulsen and Pogue (1988).\textsuperscript{13} However, in tandem with the rest of this section, the regression analysis does support the conclusion that the increase in college premiums cannot be largely explained by shifting patterns in undergraduate major choices.

The Trend in Economic Value

At the beginning of this chapter, I posited that there are two differences between college premiums and the economic value of higher education: the selection bias and the effect of undergraduate major choices. Given the increase in college premiums between 1975 and 2014, I have attempted to show that the economic value of going to college has also increased by eliminating the possibility that the increase in premiums was primarily caused by changes in the selection bias or shifts in major choices. Though the analyses in this chapter suffer from certain

\textsuperscript{13} Paulsen and Pogue (1988) show how major choices may be influenced by labor market conditions. In addition, Rumberger (1984) finds that college premiums may be positively correlated with macroeconomic growth. To the extent that both these conclusions were accurate, this could cause a simultaneity issue with the regressions in this section.
drawbacks, it seems highly likely that the economic value of possessing an undergraduate degree has in fact increased over the period under study.
Chapter 4: Conclusions and Discussion

Summary of Findings

The primary objective of this thesis is to analyze trends in college premiums between 1971 and 2014 and determine whether they varied significantly among certain demographic groups. In this thesis, college premiums are defined as the mean percentage difference in income between college and high school graduates. My analyses suggest that college premiums fell during the first half of the 1970s, from around 35% in 1971 to about 26% in 1975. After 1975, college premiums began to increase, fluctuating around an upward trend for the rest of the period under study. The most rapid increase took place in the 1980s, when premiums rose from 29% to 47%. In contrast, the increase was much slower during the second half of the 1970s and in the 2000s. My estimates indicate that the college premium is currently around 55-56%. For a more in-depth analysis of these trends, see Chapter 2.
Disaggregating college premiums by gender, I find similar results for men and women in all but two periods: the first half of the 1970s and 1985-1997. In both of these periods, college premiums were significantly higher for women. Looking next at racial differences, my results indicate that college premiums were significantly higher for blacks than for whites throughout the period. The black-white gap shrank slightly over the 44 years analyzed, averaging 16 percentage points in the first half of the period and 12 percentage points in the second. I also examined college premiums for the Hispanic, Latino, and Spanish population, finding that they matched up nearly identically with the estimates for the rest of the sample. Lastly, the sample was separated into two age cohorts, above and below age 35. Regression output reveals higher college premiums for the older cohort during the 1970s, after which point premiums became higher for the younger cohort and stayed that way for the remainder of the period. These results are analyzed and discussed in detail in Chapter 2.

As noted in Chapter 1, college premiums cannot be strictly equated to the economic value of possessing a college degree. In
Chapter 3, I explore the connection between changes in college premiums and the trend in the economic value of higher education. I present evidence that the economic value of a college degree did indeed increase alongside college premiums, even when accounting for the selection bias in college premiums and the role of shifting undergraduate major choices.

The Negative Selection Hypothesis

The analyses presented in this thesis provide support for the negative selection hypothesis, which maintains that members of the demographic groups least likely to obtain college degrees tend to reap the greatest economic rewards from doing so (see Chapter 1 for an overview of the positive and negative selection hypotheses). This is consistent with the work of Brand and Xie (2010), who argue for the applicability of the negative selection hypothesis to higher education in the United States.

Turning first to gender differences in college premiums (Chapter 2, Graph 2), my estimates indicate a higher college premium for women in the majority of years between 1971 and
1997, after which point the difference becomes statistically insignificant. These results are consistent with the predictions of the negative selection hypothesis, given college enrollment trends for men and women. In 2013, women overtook men with respect to the number of individuals ages 25 and up who had attained at least a bachelor’s degree (U.S. Census Bureau, 2014). However, since the early 1990s, females recently graduated from high school have been more likely to enroll in college than their male counterparts (Digest of Education Statistics, 2012). When discussing negative selection, I find that neither date is completely appropriate. The former is imperfect because in 2013, young women had been more likely to attend college than young men for around two decades. However, the latter is also problematic because all of the estimates after 1990 are sure to include some individuals who attended college at a time when it was still more likely for young men to be enrolled. Therefore, it is unclear exactly when the negative selection hypothesis predicts that the gender difference in college premiums should have disappeared or reversed. It can be more broadly stated, however, that women’s
college premiums should be higher than men’s in the beginning and middle of the 1971-2014 period and equal to or lower than men’s toward the end, with the transition happening sometime between the early-1990s and 2013. This is indeed evidenced by the results, making the gender difference in college premiums consistent with the negative selection hypothesis.\footnote{One caveat with this line of reasoning, as noted in Chapter 1 and by Holmlund and Sund (2008), is that female students have historically outperformed their male counterparts by most academic measures. This could provide an alternate explanation for the gender disparity between 1971 and 1997.}

The gap in college premiums between whites and blacks presents additional evidence for the negative selection hypothesis (Graph 3, Chapter 2). In this case, the situation is clearer since whites were more likely than blacks to enroll in college throughout the 1971-2014 period (Digest of Education Statistics, 2012). In addition, blacks have higher college premiums than whites in every year analyzed, and the differences are statistically significant in the majority of years. This clearly substantiates the negative selection hypothesis.

In contrast to the previous two cases, the comparison between the Hispanic, Latino, and Spanish cohort and the rest of
the sample provides no evidence for either negative or positive selection (Graph 4, Chapter 2). Because there is virtually no difference between the two groups in terms of college premiums, the only way for either of the hypotheses to be clearly supported would be for the enrollment rates of the two groups to be nearly identical. This, however, is not the case: enrollment rates are lower for the Hispanic, Latino, and Spanish population in the majority of years (Digest of Education Statistics, 2012). Although the enrollment gaps are not large—hovering around 5-12 percentage points in magnitude—they are substantial enough to render this one instance in which the negative selection hypothesis is inconsistent with results.

This last case notwithstanding, my analyses do suggest that the negative selection hypothesis is valid in certain contexts. Furthermore, there is no evidence in any of my results supporting the positive selection hypothesis. The finding that the negative selection hypothesis is inconsistent with the analysis of the Hispanic, Latino, and Spanish sample may suggest that this hypothesis is too general to apply to every inter-demographic
comparison of college premiums. This is relatively easy to accept, since the likelihood of attending college is almost certainly not the only factor influencing college premiums.

That being said, the fact that there is evidence of the negative selection hypothesis in other sections of my analysis poses an interesting question for standard human capital theory. If groups that are less likely to get a college education seem to benefit more—or at least equally—from doing so, one or more of the following must be true: (1) there is something preventing the group members from pursuing a degree, (2) there is some type of benefit not accounted for, or (3) there is a lack of awareness among certain groups regarding the economic value of college. One potential explanation is that college admissions departments are biased towards groups that are more likely to pursue an undergraduate degree. This seems unlikely, though, especially given the use of Affirmative Action policies in many universities. A more probable alternative stems from the reality that high tuition rates and limited financial aid budgets restrict access to college for members of certain socioeconomic classes. The black-white
income gap, for example, could therefore explain why whites are more likely to attend college, even though college premiums are higher for blacks.

Additionally, the perceived economic benefit of obtaining a college degree is likely just one of many reasons for attending college, and the others—the expectations of parents, the prestige associated with higher education, the social aspects, etc.—may be equally motivating. Unfortunately, parameterizing benefits such as the social aspects of college and the desire to satisfy expectations is infinitely complicated, making it impossible to produce a fully comprehensive model of the benefits of going to college. For this reason, any validation of the negative selection hypothesis must be taken with a grain of salt.

College Premiums and Income Inequality

In the introduction, I mentioned Goldin and Katz’s (2007b) finding that a large portion of the increase in income inequality in the U.S. since the 1980s can be explained by increases in education-based wage differentials such as the college premium.
This seems to imply that higher education is an important component of the larger system that fosters this inequity. Although this idea undoubtedly has merit, my results suggest that it may be misleadingly simple.

The issue with this concept becomes clear when thinking about the black-white income gap. The disparity in income between blacks and whites has remained relatively constant since the 1970s, with white households earning significantly more than black households (Plumer, 2013).\(^{15}\) Given the finding that college premiums are higher for blacks, it follows that the black-white wage gap must be smaller among degree-holders. In this sense, college actually promotes income equality. The same is also true for the gender income gap. Although college premiums for men and women have been equivalent since the late-1990s, they were larger for women during the first half of the 1970s and for the majority of 1985-1997. As a result, it can be concluded that higher education served to reduce gender income inequality during the

---

\(^{15}\) Plumer’s (2013) estimates show that black households earned only 59% as much as white, non-Hispanic households as recently as 2010.
late 20th century by bridging the gender wage gap among college graduates.

With these observations in mind, the idea that college premiums are responsible for growing income inequality appears inadequate. Although the upward trend in college premiums does indicate some degree of polarization in the wage structure, higher education also seems to be an especially effective means of achieving upward mobility for certain groups that are associated with lower median incomes. In this way, higher education actually encourages a more equitable distribution of wealth.

**Future Research**

There are numerous possible directions for future research on the topics presented in this thesis. One way of advancing this line of work would be to explore income differentials between more than two education groups. As associate degrees and graduate programs have become more and more popular, the CPS March Supplement samples have started to include large subsamples of people with these degrees. Estimating the
“community college premium” and the “master’s degree premium” would certainly add value to this discussion.\footnote{Using CPS data, this would unfortunately only be possible after 1992. Before this time, postsecondary educational attainment was measured by the number of years spent at postsecondary institutions, as opposed to the highest degree earned, making it impossible to reliably distinguish between different postsecondary degrees.}

Perhaps the most important question, on which much work has been done without arriving at a clear consensus, is why the economic value of higher education is rising. Should it be attributed to technological change fueling demand for highly skilled labor, or to a relative scarcity in the supply of college graduates? How do factors like immigration, tax policy, and unionization influence the economic value of a degree? The ultimate goal in this area of research is to integrate the array of competing theories into a comprehensive model of college premiums over time. If accomplished, this model would greatly further our understanding of the economic value of higher education and its relationship to income inequality in the United States.
References


"Percentage of Population 25 Years and Older, and 25-29 Years Old, with Bachelor’s Degree or Higher by Sex: 1947-2014." 


<table>
<thead>
<tr>
<th>Year</th>
<th>Full Sample</th>
<th>Men</th>
<th>Women</th>
<th>White</th>
<th>Black</th>
<th>Hispanic, Latino, or Spanish Origin</th>
<th>Non-Hispanic, -Latino, or -Spanish Origin</th>
</tr>
</thead>
<tbody>
<tr>
<td>1979</td>
<td>1980</td>
<td>1981</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### Note:
Shaded regions indicate statistically significant differences.
<table>
<thead>
<tr>
<th>Year</th>
<th>Age ≤ 35</th>
<th>Age &gt; 35</th>
<th>Men</th>
<th>Women</th>
<th>White</th>
<th>Black</th>
<th>Non-Hispanic, Latino, or Spanish-Origin</th>
<th>Hispanic, Latino, or Spanish-Origin</th>
<th>Full Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>1982</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1983</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1984</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1985</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1986</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1987</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1988</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1989</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1990</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1991</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1992</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: Shaded regions indicate statistically significant differences.

<table>
<thead>
<tr>
<th>Year</th>
<th>Full Sample</th>
<th>Men</th>
<th>Women</th>
<th>Age ≤ 35</th>
<th>Age ≥ 35</th>
</tr>
</thead>
<tbody>
<tr>
<td>1993</td>
<td>47%</td>
<td>48%</td>
<td>47%</td>
<td>53%</td>
<td>42%</td>
</tr>
<tr>
<td>1994</td>
<td>47%</td>
<td>46%</td>
<td>47%</td>
<td>57%</td>
<td>42%</td>
</tr>
<tr>
<td>1995</td>
<td>46%</td>
<td>46%</td>
<td>46%</td>
<td>51%</td>
<td>45%</td>
</tr>
<tr>
<td>1996</td>
<td>46%</td>
<td>46%</td>
<td>46%</td>
<td>50%</td>
<td>44%</td>
</tr>
<tr>
<td>1997</td>
<td>49%</td>
<td>48%</td>
<td>48%</td>
<td>51%</td>
<td>47%</td>
</tr>
<tr>
<td>1998</td>
<td>46%</td>
<td>46%</td>
<td>46%</td>
<td>52%</td>
<td>44%</td>
</tr>
<tr>
<td>1999</td>
<td>47%</td>
<td>47%</td>
<td>47%</td>
<td>51%</td>
<td>45%</td>
</tr>
<tr>
<td>2000</td>
<td>52%</td>
<td>51%</td>
<td>52%</td>
<td>52%</td>
<td>49%</td>
</tr>
<tr>
<td>2001</td>
<td>52%</td>
<td>51%</td>
<td>52%</td>
<td>53%</td>
<td>49%</td>
</tr>
<tr>
<td>2002</td>
<td>51%</td>
<td>52%</td>
<td>51%</td>
<td>52%</td>
<td>49%</td>
</tr>
<tr>
<td>2003</td>
<td>51%</td>
<td>52%</td>
<td>51%</td>
<td>52%</td>
<td>49%</td>
</tr>
</tbody>
</table>

Note: Shaded regions indicate statistically significant differences.

- **Hispanic, Latino, or Spanish Origin**: Non-Hispanic
- **Age > 35**: Age ≤ 35
## College Premium Estimates, by Cohort (2004-2014)

<table>
<thead>
<tr>
<th>Year</th>
<th>Full Sample</th>
<th>Men</th>
<th>Women</th>
<th>White</th>
<th>Black</th>
<th>Hispanic, Latino, or Spanish Origin</th>
<th>Non-Hispanic, -Latino, or -Spanish Origin</th>
</tr>
</thead>
<tbody>
<tr>
<td>2004</td>
<td>55%</td>
<td>53%</td>
<td>56%</td>
<td>46%</td>
<td>62%</td>
<td>44%</td>
<td>49%</td>
</tr>
<tr>
<td>2005</td>
<td>50%</td>
<td>53%</td>
<td>57%</td>
<td>45%</td>
<td>57%</td>
<td>45%</td>
<td>49%</td>
</tr>
<tr>
<td>2006</td>
<td>55%</td>
<td>51%</td>
<td>57%</td>
<td>50%</td>
<td>55%</td>
<td>49%</td>
<td>49%</td>
</tr>
<tr>
<td>2007</td>
<td>55%</td>
<td>51%</td>
<td>57%</td>
<td>50%</td>
<td>55%</td>
<td>49%</td>
<td>49%</td>
</tr>
<tr>
<td>2008</td>
<td>55%</td>
<td>51%</td>
<td>57%</td>
<td>50%</td>
<td>55%</td>
<td>49%</td>
<td>49%</td>
</tr>
<tr>
<td>2009</td>
<td>55%</td>
<td>51%</td>
<td>57%</td>
<td>50%</td>
<td>55%</td>
<td>49%</td>
<td>49%</td>
</tr>
<tr>
<td>2010</td>
<td>55%</td>
<td>51%</td>
<td>57%</td>
<td>50%</td>
<td>55%</td>
<td>49%</td>
<td>49%</td>
</tr>
<tr>
<td>2011</td>
<td>55%</td>
<td>51%</td>
<td>57%</td>
<td>50%</td>
<td>55%</td>
<td>49%</td>
<td>49%</td>
</tr>
<tr>
<td>2012</td>
<td>55%</td>
<td>51%</td>
<td>57%</td>
<td>50%</td>
<td>55%</td>
<td>49%</td>
<td>49%</td>
</tr>
<tr>
<td>2013</td>
<td>55%</td>
<td>51%</td>
<td>57%</td>
<td>50%</td>
<td>55%</td>
<td>49%</td>
<td>49%</td>
</tr>
<tr>
<td>2014</td>
<td>55%</td>
<td>51%</td>
<td>57%</td>
<td>50%</td>
<td>55%</td>
<td>49%</td>
<td>49%</td>
</tr>
</tbody>
</table>

*Note: Shaded regions indicate statistically significant differences.*