UNDERSTANDING THE EFFECT OF COMPUTER SCIENCE COURSEWORK ON GRADUATE SCHOOL ATTENDANCE

Parsa Keyvani

Submitted in partial fulfillment of the requirements

for the degree of Bachelor of Science in

Computer Science and Economics

UNION COLLEGE

2023

Acknowledgements

I would like to express my deepest gratitude and appreciation to Prof. Stephen Schmidt and Prof. Aaron Cass for their invaluable guidance and support throughout the journey of completing this thesis. Their expertise, encouragement, and commitment to excellence have been instrumental in shaping the direction and quality of this research. This thesis would not have been possible without the collective support and guidance from all these individuals. I am truly grateful for their contributions, and I acknowledge their invaluable role in shaping the outcome of this research.

Abstract

UNDERSTANDING THE EFFECT OF COMPUTER SCIENCE COURSEWORK ON GRADUATE SCHOOL ATTENDANCE

Parsa Keyvani

This study investigates the impact of undergraduate computer science coursework on graduate school attendance and post-college income. Using data from the National Center for Education Statistics Baccalaureate and Beyond survey, the relationship between computer science coursework, graduate school attendance, and post-college income is examined through logistic and linear regression analyses. The findings reveal that completing computer science coursework positively influences the likelihood of attending graduate school, highlighting the importance of computer science education in fostering interest and motivation for advanced studies. However, the linear regression analysis does not demonstrate a significant association between computer science coursework and post-college income. Future research utilizing alternative models, such as exponential models, may provide further insights into this relationship. These findings have implications for educational institutions and policymakers in enhancing computer science programs and promoting career opportunities in the field.

Table of Contents

CHAPTER 1: INTRODUCTION

1.1 Statement of the Problem

Post-secondary education plays a crucial role in shaping individuals' future prospects, with graduate education serving as a pathway to advanced knowledge and enhanced career opportunities (Bailey & Dynarski, 2011; Carnevale, Smith, & Strohl, 2010). While there is existing research on the determinants of graduate school enrollment, such as socioeconomic background, academic performance, and field of study (Mullen, Goyette, & Soares, 2003; Posselt & Grodsky, 2017), there is a significant gap in our understanding of the specific impact of undergraduate computer science coursework on individuals' likelihood of pursuing graduate studies. Although some studies have examined the effect of majoring in computer science on graduate school attendance, the influence of specific coursework within the discipline remains largely unexplored. Thus, this research aims to address this gap by focusing specifically on the effect of computer science coursework on graduate school attendance, providing valuable insights into the unique contribution of coursework in computer science to individuals' decision to pursue advanced education.

Understanding the impact of computer science undergraduate courses on graduate school enrollment holds significant importance for multiple reasons. Firstly, in the face of increasing demand for highly skilled professionals in technology-related fields, it is crucial to identify the factors that influence individuals' decisions to pursue further education in computer science. Secondly, as the job market becomes more competitive, individuals with advanced degrees often enjoy enhanced opportunities and higher earning potential. However, it is worth noting that there is also a demand for individuals with undergraduate degrees in computer science alone.

Moreover, investigating the relationship between computer science undergraduate education and graduate school enrollment can provide valuable insights for curriculum design and program offerings in educational institutions. If it is found that computer science courses positively impact the likelihood of attending graduate school, academic institutions may consider enhancing their computer science programs to better prepare students for advanced studies. On the other hand, if such courses are found to hinder graduate school attendance, interventions and support structures can be developed to address potential barriers or misconceptions that deter students from pursuing further education. It is important to note that graduate school enrollment extends beyond STEM fields, and differentiating between STEM and non-STEM programs would provide a more comprehensive understanding of the impact of computer science coursework on graduate school attendance.

1.4 Research Questions

This study explores the following questions

1) How does the number of Computer Science courses completed at the undergraduate level impact the probability of attending graduate school?
2) What is the association between undergraduate Computer Science coursework and post-college income levels?

3) Do factors such as gender, race/ethnicity, GPA, major, citizenship status, and family income mediate or moderate the relationship between Computer Science coursework and post-college income?

1.3 Differences from Previous Studies

This study distinguishes itself from previous research in the field by focusing specifically on the relationship between undergraduate computer science coursework and graduate school attendance. While numerous studies have explored the determinants of graduate school enrollment, particularly in relation to factors such as race, gender, socio-economic status, and academic performance (GPA), few studies have specifically examined the influence of academic factors, such as the effect of STEM enrichment programs and graduate school attendance. Thus, this research addresses a significant gap in the existing literature by delving into the unique relationship between undergraduate computer science coursework and the likelihood of attending graduate school.

Furthermore, while previous studies in the field, such as Perna (2004) and Peters & Daly (2013), have relied on older datasets and focused on narrow timespans when examining enrollment behavior, this study uses a more recent dataset. By using data from the National Center for Education Statistics Baccalaureate and Beyond survey, covering the period from 2008 to 2018, this research benefits from a contemporary understanding of the factors influencing graduate school attendance. Although the dataset primarily includes individuals who obtained their undergraduate degrees around 2008, the study takes into account their post-college activities over a longer period, providing insights into their decisions regarding graduate school attendance. The inclusion of this dataset sets this study apart from previous research and offers a more accurate reflection of the current landscape of graduate school enrollment and its relationship with computer science undergraduate coursework.

Additionally, this study expands upon existing models developed by researchers such as Merolla and Serpe (2013) and Xu (2014) by including a broader range of variables based on a comprehensive review of the literature. By incorporating factors such as gender, race/ethnicity, GPA, major, citizenship status, socio-economic status, and family background, this research presents a more nuanced analysis of the determinants of graduate school enrollment. The inclusion of these variables enhances the explanatory power of the model and provides a more comprehensive understanding of the relationship between undergraduate Computer Science courses and graduate school attendance.

1.5 Thesis Structure

The thesis is structured into seven chapters. Chapter 1 introduces the study and presents the problem statement, objectives, and research questions. Chapter 2 serves as the theoretical background, providing a foundation for understanding the rise of graduate school attendance and computer science education, as well as the impact of CS coursework on graduate school attendance. In Chapter 3, a comprehensive review of the relevant literature is presented, examining previous research on the determinants of graduate school enrollment and the cyclical nature of enrollment. Building upon the theoretical and literature review, Chapter 4 outlines the hypotheses formulated for the study. Chapter 5 describes the data sources and methods used in the research, including the National Center for Education Statistics Baccalaureate and Beyond survey. Chapter 6 presents the results obtained from the linear and logistic regression analyses, analyzing the relationship between undergraduate computer science coursework, graduate school attendance, and post-college income. Finally, Chapter 7 provides the conclusion, summarizes the key findings of the study, and offers recommendations based on the research outcomes.

CHAPTER 2: BACKGROUND ON EDUCATION

2.1 The Growth of Graduate Education

Graduate degrees in STEM offer numerous opportunities and can significantly advance one's career. In recent decades, the value and importance of graduate education have increased substantially (Buttet and Schoonbroodt 2022, 333-71). This growth can be attributed to the increasing demand for and supply of graduate education (Blagg, 2018).

2.1.1 The Demand Side Factors

On the demand side, several factors have led to the growth of graduate education. First, due to the rapid advancement of information and technology, a college education is no longer sufficient for many occupations (Schneider and Klor de Alva, 2018; Lynch, Grogan & Willems, 2016). As a result, graduate degrees have become prerequisites for most top-ranked jobs in fields such as business, education, health care, and STEM (Schneider and Klor de Alva, 2018; Lynch, Grogan & Willems, 2016; Carnevale et al., 2016). Second, pursuing a graduate degree is an investment in one's future, as it can lead to higher salaries and better career opportunities (Blagg, 2018; Blagg, 2022; Glazer-Raymo, 2005). According to the Bureau of Labor Statistics, workers holding a master's degree earn an average of $12,000 more annually than those with only a bachelor's degree, with the premium varying substantially by occupation ("Education Pays," 2022).

The third factor contributing to the growth of graduate education is the employers' perspective, who view master's programs as a way to promote employee advancement and retention (Blagg, 2018; Bedard and Herman, 2008). Many employers are now willing to subsidize some or all of their employees' tuition fees because they recognize the value of a master's degree in helping employees gain new skills that can be applied to their job (Blagg, 2018). Additionally, obtaining a master's degree can lead to higher levels of job satisfaction and employee engagement, leading to increased productivity and profitability (Lynch, Grogan & Willems, 2016; Blagg, 2018).

The increasing number of international students in the U.S. is another factor contributing to the growth of graduate education (Chen and Fink, 2019; Blagg, 2018). The share of international students enrolled in American master's programs has risen from 4 percent in 1996 to 13 percent in 2016 (Chen and Fink, 2019). For many of these students, obtaining a graduate degree offers an opportunity to remain in the country for an extended period. International students often face difficulties in obtaining work visas after completing their undergraduate studies, which makes pursuing a graduate degree an attractive option as it can extend their stay in the U.S.

2.1.2 The Supply Side Factors

There are also factors on the supply side that contribute to the growth of graduate education. First, graduate programs, specifically master's programs, have grown more diverse in recent years. Master's programs have gradually enrolled a larger share of students from underrepresented racial and ethnic backgrounds. The share of black and Hispanic students enrolled in master's programs has nearly doubled in 20 years, from 14 percent in 1996 to 25 percent in 2016. Similarly, the share of white students has declined, from 84 percent of all master's students in 1996 to 57 percent in 2016 (Baum and Steele, 2017; Blagg, 2018).

Just as master's students have changed, so have the programs that offer these degrees. From 1995 to 2017, the number of distinct master's degree program fields has grown substantially (Schneider and Klor de Alva, 2018). The number of distinct master's fields granted at least 100 degrees nationally yearly has risen from 289 to 514 (Schneider and Klor de Alva, 2018). This trend is evident in the field of computer science (where at least one master's degree was granted in only seven unique fields in 1995 and in 25 fields in 2017) and in the field of biology (where 31 unique master's program fields in 1995 became 81 in 2017) (Schneider and Klor de Alva, 2018; Blagg, 2018). However, this specialization has also occurred in fields with less scientific or technological change, such as in education and visual and performing arts (Blagg, 2018; AIR, 2013).

Unlike undergraduate institutions, master's programs are not constrained by the same limitations and pressure to increase enrollment. For instance, a significant percentage of first-year bachelor's degree seekers reside on campus, which means expanding the size of the freshman class could require expanding dormitory space, cafeterias, and other facilities, which can be costly and challenging for institutions to manage (Baum and Steele, 2017; AIR 2013). Additionally, institutions may face pressure from surrounding neighborhoods to limit the number of on-campus undergraduates. In contrast, most master's students live off-campus and commute to school, which means they require less dormitory space and may be more flexible in attending classes held in the evening or on weekends. Also, increasing undergraduate enrollment may lead to a decrease in selectivity, which is often seen as a measure of institutional quality (Carnevale, Cheah, & Hanson, 2015). In contrast, master's programs may not face the same pressure to maintain selectivity and competitiveness, as acceptance rates for these programs are often higher than those for bachelor's degree programs. This disparity in acceptance rates indicates that selectivity, as traditionally measured, may not be the primary factor in evaluating the prestige of master's programs. Other factors like faculty expertise, research opportunities, and industry connections may hold more importance in assessing the quality and prestige of master's programs. (AIR 2013).

Another key factor driving the growth of graduate education is the increasing availability of online programs (Goodman, Melkers, & Pallais, 2018; Burd, Smith, & Reisman, 2015; Hoxby, 2014). This trend is particularly attractive to self-directed learners who wish to advance their education while continuing to work. Over the years, the percentage of master's students enrolled in fully online programs has risen dramatically. In 2004, only 9 percent of part-time and 5 percent of full-time students reported enrolling exclusively in distance education. However, by 2016, those figures had risen to 36 percent and 27 percent, respectively (Burd, Smith, & Reisman, 2015). Notably, online enrollment rates for master's programs far outpace those for bachelor's degree programs, with 31 percent of master's students reporting that their program was fully online in 2016, compared to just 12 percent of bachelor's students (Hoxby, 2014).

2.2 The Rise of Computer Science Education

Computer science education has experienced a significant surge in recent years, driven by a combination of demand-side and supply-side factors. On the demand side, factors such as technological advancements and the abundance of career opportunities have led to an increased interest in computer science education. It is important to note that this encompasses any level of computer science education, including individual courses or a major in computer science. On the supply side, the expansion of educational offerings and government initiatives has played a crucial role in meeting the growing demand for computer science education.

2.2.1 The Demand Side Factors

The rapid advancement of technology has transformed numerous industries and created a high demand for individuals with computer science skills. Emerging technologies such as artificial intelligence, machine learning, and big data require a workforce equipped with the knowledge and expertise to develop, implement, and manage these technologies. As a result, there is a growing demand for computer science education to cultivate the necessary skills to meet these technological challenges.

Computer science offers a wide range of lucrative career opportunities (Dalton and Goodrum 1991, 483-506). In today's digital era, almost every industry relies on technology and requires professionals with computer science expertise. Fields such as software development, data science, cybersecurity, computer networking, and web development have experienced significant growth and offer well-paying jobs. Individuals who develop expertise in computer and technology fields earn higher wages and employment. Even those who do not pursue technical occupations still reap these benefits, as computing and data analysis skills have been broadly integrated into many industries and occupations. Therefore, the promise of rewarding and fulfilling careers attracts students to pursue computer science education.

2.2.2 The Supply Side Factors

Educational institutions, from K-12 schools to colleges and universities, have recognized the increasing demand for computer science education. In response, they have expanded their computer science offerings to cater to student interests and industry demands. Many institutions now offer computer science courses, majors, and degree programs, providing students with the opportunity to specialize and acquire the necessary skills in this field. This expansion of educational offerings has significantly contributed to the supply of computer science education.

U.S. Government initiatives have played a significant role in driving the increase in computer science (CS) education across the country. The government acknowledges that economies overall fare better when individuals are more technologically competent (Hansen and Zerbino, 2022). Studies demonstrate a positive relationship between economic growth, technology, and investments in human capital related to CS skills (Maryska, Doucek, and Kunstova 2012). Many states and countries view computing and technology jobs as engines of economic growth; thus, providing public school students with quality CS education enables sustainable growth. Federal and local politicians often appeal to this economic rationale to justify investments in CS education to public stakeholders. Arkansas, an early adopter of CS education policies, serves as a prime example of successful implementation and the positive outcomes achieved. One of the notable achievements in Arkansas was the establishment of incentives to encourage participation in CS education programs. Since the 2015-16 school year, students have been able to use CS courses to fulfill high school science and mathematics requirements. This policy change has contributed to increased enrollment and improved diversity in CS classes. Additionally, Arkansas developed CS certification for in-service and pre-service teachers, encouraging educators to participate in training programs (Fowler and Vegas, 2022).

2.3 The Impact of CS Coursework on Graduate School Attendance

Given the growing importance of graduate education and the increasing demand for skilled professionals in computer science, it is important to examine how taking CS classes in their undergraduate program may impact an individual's decision to attend graduate school. Taking CS courses could have both positive and negative effects on the likelihood of attending graduate school. The subsequent paragraphs will delve into the factors that might contribute to a favorable or unfavorable outcome in this relationship.

2.3.1 CS coursework may increase grad school enrollment

There are three primary ways in which taking CS courses can positively impact an individual's decision to pursue graduate school. Firstly, the economic theory of intrinsic motivation and self-selection supports the notion that CS coursework can increase graduate school enrollment. Many students who choose to pursue computer science have a genuine passion for the subject and possess a strong intellectual curiosity that compels them to delve deeper into its intricacies. Undergraduate classes in computer science can act as a catalyst for nurturing and fueling these personal interests. As students engage in coursework and discover the breadth and depth of the field, their passion and curiosity intensify, driving them toward a desire for more comprehensive knowledge and understanding.

Moreover, CS coursework has the potential to increase grad school enrollment by equipping students with relevant skills that make them stand out in the applicant pool. By completing undergraduate computer science classes, students acquire a solid foundation in key concepts, programming languages, and problem-solving techniques. This proficiency not only enhances their technical abilities but also instills confidence in their capacity to handle graduate-level coursework. As they apply to graduate programs, having a demonstrated background in CS can set them apart from other applicants who may not have similar exposure or expertise. This competitive edge further reinforces their belief in their abilities, making them more inclined to pursue advanced studies.

2.3.2 CS coursework may decrease grad school enrollment

While CS coursework can enhance the likelihood of attending graduate school through the factors discussed earlier, it is important to recognize that there are also instances where it may have the opposite effect and potentially decrease graduate school enrollment. Several factors contribute to this possibility. First, CS coursework may decrease grad school enrollment due to the high demand for professionals with computer science skills in various industries. Upon completing undergraduate computer science classes and acquiring marketable skills, individuals often find themselves presented with immediate job opportunities that align with their career aspirations. This aligns with the theory of human capital, where individuals invest in education and training to enhance their productivity and earning potential. By obtaining relevant skills through CS coursework, individuals increase their human capital and become attractive candidates in the job market (Giannakos and Jaccheri 2017). A fulfilling and well-paid position may lead some individuals to perceive a reduced need for graduate education to secure desirable employment. As a result, the motivation to pursue advanced studies may be diminished, potentially resulting in a decrease in graduate school enrollment.

Potential salary growth by taking CS courses can also contribute to a decrease in graduate school enrollment. The computer science industry is known for its competitive salaries and promising opportunities for career advancement. Individuals who have completed undergraduate computer science classes may carefully consider the potential salary growth and career progression they can achieve by entering the job market directly, without pursuing further education. They weigh the anticipated financial benefits and professional growth against the additional time and financial investment required for graduate education. If the projected salary growth and career prospects without a graduate degree appear substantial, individuals may be inclined to prioritize immediate employment and forego pursuing advanced degrees. This consideration of potential salary growth can significantly impact the decision-making process and potentially lead to a decrease in grad school enrollment among those who have completed CS coursework.

The decision to pursue graduate education involves a substantial investment of both time and financial resources. Consequently, some individuals may be deterred from attending graduate school when they weigh the opportunity cost of forgoing immediate entry into the job market. Individuals who have acquired employable skills through their undergraduate computer science classes may find themselves with enticing job prospects. Therefore, the net utility gained from entering the workforce immediately and capitalizing on these job opportunities may outweigh the perceived benefits of pursuing advanced degrees. This can contribute to a decrease in grad school enrollment among those who have completed CS coursework.

CHAPTER 3: REVIEW OF THE LITERATURE

Overview

This section synthesizes available literature on factors affecting students' decision to attend graduate school. Firstly, we introduce the contextual factors that have led to the recent increase in graduate school enrollment. We then focus our attention on the patterns of graduate school enrollment and persistence. This includes closely examining the literature on each influential factor affecting graduate school enrollment and persistence. These factors include demographic information, such as gender, race, and family socioeconomic status, as well as academic information such as the number of computer science courses taken in college, college GPA, and major (STEM vs. Non-STEM).

Students' Decision to Attend Graduate School

The existing literature on post-baccalaureate enrollment and degree attainment is notably limited, especially when compared to the extensive body of research available at the high school and undergraduate levels. While a few published studies have contributed to our understanding of post-baccalaureate school choice, they have primarily focused on specific graduate degree programs, institutions, or have utilized outdated data. The higher education has undergone significant transformations over the past 15 to 20 years, marked by shrinking governmental subsidies (Barr & Turner, 2013), rise in tuition costs, and a growing reliance on student debt (Dynarski & Scott-Clayton, 2013; Carlton, 2015; English & Umbach, 2016). These factors have further complicated the dynamics of post-baccalaureate education (Dynarski & Scott-Clayton, 2013; Morreale & Staley, 2016).

Determinants of Graduate School Enrollment

Numerous empirical studies investigating the determinants of graduate school enrollment have employed sociological theories to examine the factors influencing students' enrollment decisions. Researchers have explored the impact of student and family background characteristics on post-baccalaureate educational attainment and economic success, highlighting the influence of social and cultural factors. These factors include parents' education, family income, undergraduate field of study, undergraduate institutional characteristics, and undergraduate debt, which are believed to have significant effects on students' educational and career trajectories.

2.1.1 Family Background and Academic Experience

Family background characteristics, such as first-generation status, parental education, and family income, have been shown to have a direct impact on various aspects of undergraduate education. Research indicates that these factors influence the type of undergraduate institution students attend, their level of integration within the university, as well as their academic performance, expectations, and career values (Ethington & Smart, 1986; Mullen, Goyette, & Soares, 2003; Zhang L., 2005). First-generation students, whose parents often have lower levels of education, exhibit lower aspirations for and lower rates of attendance in graduate school compared to students with more educated parents (Carlton, 2015). Moreover, when first-generation students do pursue graduate education, they are less likely to complete their degrees (Mullen, Goyette, & Soares, 2003). These influences during the undergraduate years subsequently impact the decision to pursue graduate education indirectly.

Previous studies have utilized different datasets to investigate the influence of background characteristics on graduate enrollment. For instance, Ethington and Smart (1986) employed the Cooperative Institutional Research Program, a survey conducted on the entering freshman class in 1971 and followed up in 1980, to examine graduate enrollment. Paulsen and St. John (2002) utilized the National Postsecondary Study Aid Survey of 1987 (NPSAS87) for their analysis. Mullen et al. (2003), Zhang (2005), and Perna (2004) utilized the Baccalaureate and Beyond Longitudinal Survey (BB:93/97), while English and Umbach (2016) conducted their analysis using the B&B:2000/01 dataset.

The findings of Ethington and Smart (1986) suggest that, after controlling for field of study and financial aid, social and academic integration had the strongest influence on the decision to enroll in graduate school. However, their analysis focused on overall enrollment without delving into different types of graduate degree programs. Other researchers have further disaggregated graduate education, examining separate degree programs, and have found that considering different program types provides more insights into the impact of family background. For example, Mullen et al. (2003) discovered that parents' education had a significant effect on enrollment in professional and doctoral programs, a smaller effect on master's programs, and no effect on MBA programs. However, once test scores and undergraduate institutional characteristics were accounted for, the impact of parental education diminished, suggesting that the influence of parents' educational background varies across different program types.

Researchers have also examined the influence of undergraduate institution type and quality on persistence in graduate education (Mullen, Goyette, & Soares, 2003; Perna, 2004; Zhang L., 2005; English & Umbach, 2016). These studies have consistently found that the type of undergraduate institution attended is a significant determinant of graduate school enrollment, particularly for doctoral programs at research universities (Zhang L., 2005; English & Umbach, 2016). Mullen et al. (2003) introduced a control measure for selectivity by considering average SAT scores. They discovered that the likelihood of pursuing a graduate degree increases significantly for students who attended public and private research universities, as well as selective institutions. These findings highlight the importance of undergraduate institution type and selectivity in influencing students' decisions to pursue advanced degrees.

Furthermore, Zhang (2005) discovered that academic achievement is a strong predictor of graduate degree enrollment and attainment. In the analysis, the author examined the impact of undergraduate institutional quality on graduate school enrollment, separately considering master's and doctoral programs. Binomial logistic regression was utilized to analyze the outcomes of graduate enrollment (yes/no), type of program (master's or doctoral), and attainment of a graduate degree (yes/no). Multinomial regression was employed to examine the type of graduate institution attended (comprehensive, doctoral, or research institution). The analysis controlled for selection bias by including demographic, race/ethnicity, academic, and family background characteristics such as undergraduate major, GPA, family income, and whether the student was a first-generation college graduate. Zhang's findings indicate that first-generation college students were less likely to enroll in graduate programs, with a decrease of 2.8% and 2.15% for master's and doctoral programs, respectively. Family income had a small positive effect on master's level enrollment, but a small negative effect on doctoral enrollment. The reason for the negative effect of family income on doctoral enrollment is not explicitly clear. One possible explanation could be that doctoral programs often have lower direct costs due to the availability of scholarships and fellowship opportunities. Alternatively, it is possible that students from high-income families are more inclined to pursue other types of degrees, such as professional degrees in medicine, business, or law.

Perna (2004) utilizes a multinomial logistic regression model to investigate gender and racial disparities in graduate enrollment. The study incorporates measures of cultural capital, such as parental educational attainment and language spoken at home, as well as indicators of social capital related to the undergraduate institution, including family support, institution type, tuition, and whether the institution is out of state. The analysis reveals that, after controlling for race and ethnicity, women display a higher likelihood of enrolling in graduate degree programs compared to men. However, they are less likely to enroll in professional degree programs. The findings further demonstrate that lower starting salaries and higher GPAs significantly influence master's degree enrollment, with the effect being more pronounced for women than men. Perna concludes that the factors shaping post-baccalaureate enrollment may vary based on the student's gender.

Building upon the research conducted by Perna (2004), English and Umbach (2016) extend the investigation by utilizing a more recent survey, the B&B:2000/01, and employing a generalized hierarchical linear model that considers students clustered within their respective undergraduate institutions. The authors specifically examine students' aspirations, application processes, and enrollment decisions regarding graduate degree programs. Consistent with previous findings, they observe that both the choice of major and undergraduate GPA significantly influence students' aspirations, application behavior, and enrollment into graduate programs. However, in contrast to previous research results, English and Umbach do not find a statistically significant impact of student debt on the graduate choice process. Additionally, they do not identify significant differences based on gender or race/ethnicity. The authors conclude that minorities do not appear to face inherent disadvantages and suggest that students' experiences at their undergraduate institutions may enhance their social and/or cultural capital, enabling them to overcome inherited social and cultural constraints.

Overall, an individual's family background and academic experiences are influential in shaping their decision to pursue graduate school. The previous studies suggest that incorporating these factors into our model provides a more comprehensive explanation of post-baccalaureate enrollment decisions.

2.1.2 Undergraduate Field of Study

There is a strong correlation between the undergraduate field of study and the type of graduate degree pursued (Altonji, Arcidiacono, & Maurel, 2016). Based on the B&B:93/97 dataset, Mullen et al. (2003) and Zhang (2005) discovered that students majoring in biology and math sciences exhibited the highest likelihood of enrolling in doctoral programs, while business majors had the lowest probability of pursuing master's or doctoral degrees within four years of graduating college. Furthermore, respondents aspiring for financial gains had twice the odds of entering an MBA program, while rating career success as an important factor had a negative impact on master's degree enrollment.

However, It is important to note a limitation that restricts the generalizability of these studies: the data primarily captures individuals early in their careers. The time frame of data collection does not sufficiently cover the typical age at which individuals earn graduate, professional, and doctoral degrees. Data used by Mullen et al. (2003) and Zhang (2005) revealed that less than 30% of the students had enrolled in a post-baccalaureate program at the time of the survey. It is important to note that the motivations for students to pursue graduate education shortly after completing their undergraduate degrees may differ from those later in their careers.

2.1.3 Impact of Undergraduate Debt

Several studies have explored the negative impact of educational debt on the inclination to pursue graduate education. Research conducted by Millet (2003), Perna (2004), Carlton (2015), has demonstrated the significance of financial resources in students' decision-making regarding graduate degree programs. Students burdened with substantial undergraduate debt are generally less inclined to apply for graduate programs compared to their peers who have lower levels of educational debt.

To better understand the decision-making process of minorities in STEM fields regarding graduate school enrollment, Malcom and Dowd (2012) conducted an analysis using a constrained multinomial probit regression model with propensity score matching (PSM). The authors divided the group of borrowers into two subsets: heavy borrowers and typical borrowers, and compared them to a control group of non-borrowers. Utilizing data from the 2003 National Survey of Recent College Graduates (NSRCG), they observed considerable variation in borrowing patterns across different racial groups as shown in Table 2.

Table 2 – Level of Borrowing by Race

	Did Not Borrow	Borrowed	
		Typically	Heavily
African American	20%	48%	32%
Asian	40%	47%	13%
Latinos	28%	51%	21%
White	36%	43%	22%

Source: Author's tabulation of results from Malcom & Dowd (2012)

The study revealed that undergraduate debt appears to hinder post-baccalaureate educational opportunities, as heavy borrowing had a statistically significant and negative impact on graduate school enrollment for Latinos and Whites, but did not yield a significant effect for African Americans or Asians.

2.1.4 Impact of Recent Governmental Policies

Several recent governmental initiatives have been implemented to promote post-secondary education, including raising federal limits on student borrowing, introducing federal loan forgiveness programs, and establishing state-funded merit-based aid programs. The Higher Education Reconciliation Act of 2005 (HERA 2005) played a significant role in this regard by removing the cap on graduate school debt through the Grad PLUS loan. Starting in 2006, this program provided funding up to the total cost of attendance, replacing the previous limit of $20,500.15 Although the evidence suggests that increasing the borrowing limit has led to a rise in graduate school enrollment, no causal studies have yet examined these relationships. Additionally, the College Cost Reduction and Access Act of 2007, implemented by the Federal Government, introduced the Public Service Loan Forgiveness (PSLF) program to relieve the burden of educational debt. PSLF offers debt relief to students working in the public sector or non-profit organizations, with complete loan forgiveness granted after 120 monthly payments for eligible employees. Delisle (2016) found that 80% of PSLF borrowers utilized the program to finance graduate and professional education. Moreover, a 2018 survey conducted by the Association of American Medical Colleges Medical School Graduation Questionnaire revealed that 45.7% of medical school graduates plan to utilize a loan forgiveness program, compared to 39.7% in 2014 and 11% in 2010. It has been observed that merit- and need-based aid programs are more effective and cost-efficient in encouraging lower-income students to complete their undergraduate education and pursue graduate studies, resulting in increased probability of earning a graduate degree and higher earnings.

Impact of Business Cycle on Graduate Enrollment

Previous research has demonstrated the significant impact of economic volatility, especially on individuals entering the labor market, such as recent college graduates (Blom, Cadena, & Keys, 2015). Therefore, in order to gain insights into future patterns of graduate enrollment, it is crucial to broaden our analysis by examining literature on labor market dynamics and overall macroeconomic conditions that can influence the demand for and completion of graduate education.

According to Becker's human capital model (Becker, 1964), economic downturns lead to reduced labor demand and wages.. However, studies examining the impact of labor market conditions on college enrollment, without distinguishing between undergraduate and graduate education, have yielded conflicting findings. Some studies indicate that during a recession, students are more likely to pursue post-secondary education and continue their studies (Sakellaris & Spilimbergo, 2000; Christian, 2007; Long, Goldhaber, & Huntington-Klein, 2015), while others suggest that poor labor market conditions drive students to pursue degrees that directly lead to careers without the need for further schooling (Blom, Cadena, & Keys, 2015). The discrepancies in these findings can be attributed to variations in the data sets, cohorts analyzed, and the specific variables of interest considered.

Bedard and Herman (2008) conducted a study investigating the influence of labor market fluctuations on graduate school enrollment. Using the National Survey of Recent College Graduates (NSRCG) data from 1993 to 2001, the authors employed a latent variable model to estimate the probability of science and engineering majors enrolling in post-baccalaureate programs 19 months after obtaining their bachelor's degree. Their model considered a two-period planning horizon where students could either enroll in graduate school in period one and work in period two or work in both periods. State unemployment rates were used as substitute for the business cycle, and wages in periods one and two were captured for graduate degree holders and non-holders. The authors came up with the following latent variable model to describe the choice of enrolling in a graduate program within one to three years after completing a bachelor's degree:

$$G^* = \beta_1 UER + \beta_2 GPA + \beta_3 MAJ + \beta_4 X + \varepsilon$$

In the equation, G* represents the propensity of an individual to enroll in a post-baccalaureate program directly after college, UER denotes the state-level unemployment rate, GPA is a set of three GPA categorical variables, MAJ consists of dummies indicating the student's undergraduate major field of study, and X encompasses additional individual characteristics, graduation year, and state indicators. The authors estimated the above equation using a probit model separately for enrolling in Master's, Professional, and Doctoral degrees. State unemployment rates were assigned to individuals based on the year and state of their undergraduate education, and the analysis controlled for individuals earning their Bachelor's degree outside their home state. The findings revealed that the decision to pursue a graduate degree was influenced by gender, degree type, and GPA. Notably, women's enrollment in all graduate degrees was found to be acyclical, male master's degree enrollment was procyclical, and enrollment in doctoral programs was countercyclical for males with high GPA as well as for males with undergraduate degrees in STEM fields. Interestingly, similar results were observed when using the state unemployment rate for the student's home state, suggesting that home states may share similar unemployment rates with the states where institutions are located or that undergraduate enrollment tends to come from nearby neighborhoods with comparable unemployment rates.

In conclusion, while the existing literature provides insights into the factors influencing graduate enrollment, including gender, race, debt, undergraduate institution type, ability, and social and cultural capital, many studies are outdated or limited in scope. Changes in higher education over the past two decades, such as reduced governmental subsidies, increased tuition costs, and a more diverse student population with different attitudes and expectations, warrant further investigation. Moreover, the impact of economic conditions on graduate school enrollment varies significantly based on race, gender, program of study, and degree level. The most recent studies utilize the B&B:00/01 survey, while others focus on specific fields of study or institutions. Although these studies have advanced research in the field, their specific findings hinder generalizations to the broader context of post-baccalaureate education.

CHAPTER 4: RESEARCH DESIGN

This chapter delves into the research design employed to investigate the determinants that influence students' decision to pursue graduate education, specifically within the context of undergraduate computer science (CS) courses. Building upon the literature and trends discussed in the previous chapters, the objective is to develop a comprehensive model of college choice that incorporates the dynamics of graduate school enrollment.

The theoretical model underlying this study acknowledges that the decision to pursue graduate education involves a complex interplay of various factors beyond a straightforward assessment of monetary costs and benefits. It posits that empirical patterns of graduate enrollment are influenced by a combination of both observable and unobservable social and cultural factors, in addition to the individual characteristics typically examined in economic studies of school choice. These individual characteristics, such as age, gender, race and ethnicity, and family income, play a critical role in shaping the decision to pursue graduate education by acting as factors that can enhance or constrain the perceived costs and benefits. It is important to note that the concept of costs and benefits extends beyond purely monetary considerations, encompassing a broader range of factors that influence the decision-making process.

This chapter is structured as follows: first, an overview of the data used for the analysis is provided. Then, the key variables included in the model are defined and discussed. Next, descriptive statistics are presented to provide a comprehensive understanding of the sample characteristics. Subsequently, the methodology used to estimate the linear probability model is described. The results of the regression analysis are then presented and interpreted. Finally, the limitations of the study are discussed, recognizing the constraints and potential implications of the research design.

4.1 Data

The data used in this study was derived from the National Center for Education Statistics Baccalaureate and Beyond Longitudinal Study (B&B 2008-2018). The B&B study tracks a cohort of students who earned their bachelor's degrees in the 2007-08 academic year. The study utilizes follow-up surveys conducted at one year, four years, and ten years after graduation to collect data on various aspects of the participants' educational and professional journeys. The B&B survey encompasses information on undergraduate experiences, demographic backgrounds, labor participation, and relevant variables for this study, such as expectations regarding pursuit of graduate education, income, and debt repayment ability.

It is worth noting that the analysis of the B&B:08-18 dataset was conducted using a pre-approved computer located in Schaffer Library at Union College, in compliance with data access restrictions and ethical considerations. By leveraging this rich longitudinal dataset, we can gain valuable insights into the relationship between CS coursework and graduate school attendance, while accounting for other relevant factors.

4.2 Variables

Table 1 in this analysis presents the definitions of the independent variables used in the study. This table provides a detailed description of each variable as well as their respective data types. The definitions help ensure clarity and consistency in understanding the variables employed in the analysis. Additionally, Table 2 offers valuable information by presenting the percentage distribution, mean, maximum, and minimum values for continuous and binary variables in the sample. This summary statistics table provides an overview of the distribution and central tendencies of these variables, aiding in the understanding of their characteristics within the dataset. The selection of these variables for inclusion in the model was guided by a comprehensive review of relevant research on graduate enrollment. By drawing upon existing literature, the study aimed to ensure that the chosen variables captured essential factors known to influence graduate school attendance. This approach enhances the theoretical and empirical foundations of the analysis and strengthens the interpretation of the results.

Table 1: Variable Description and Type

Label	Description	Variable Type
GRAD_SCHOOL	Attended graduate school or not (Enrolled in either professional practice doctoral degree program master's degree program, post-master's certificate program, associate's degree program, research doctoral degree program between BA completion and B&B:08/18 interview)	Binary
QECSCERN	Number of undergraduate CS credits earned	Continuous
GPA	Undergraduate grade point average	Continuous
QFMJSTEM	STEM major field of study indicator	Binary
isFemale	Is a female or not	Binary
isOtherGender	Is other gender or not	Binary
isBlack	Is Black or not	Binary
isHispanic	Is Hispanic or not	Binary
isAsian	Is Asian or not	Binary
isAmerIndian	Is American Indian or not	Binary
isNativeHawaiian	Is Native Hawaiian or not	Binary
isMixedRace	Is mixed race or not	Binary
isOther	Is other race or not	Binary
AGE	Student's age as of 12/31/2007	Continuous
B3PAREDUC	Highest education level attained by either parent as of 2018	Categorical
not_us_citizen	Is a US citizen or not	Binary
CINCOME	Total income in 2006 for independent student or parents of dependent students	Continuous
B1OWAMT1	Indicates the cumulative amount owed as of 2008-09 on loans from all sources borrowed for the respondent's undergraduate education	Continuous
isFirstGen	Is first generation college student or not	Binary
isSecondGenBothPar	Is second generation college student (Both parents foreign)	Binary
isSecondGenOnePar	Is second generation college student (One parent foreign)	Binary

Table 2: Student and Family Background Characteristics

Label	Minimum	Maximum	Mean	Percentage
GRAD_SCHOOL	0	1	0.5	53.5% (attended grad school)
QECSCERN	0	176	3.8	44.5% (had at least one CS credit)
GPA	100	400	335	-
QFMJSTEM	0	1	0.5	30.8% (are STEM major)
isFemale	0	1	0.5	58.9%
isOtherGender	0	1	0.5	0.6%
isBlack	0	1	0.5	8.7%
isHispanic	0	1	0.5	9%
isAsian	0	1	0.5	5.8%
isAmerIndian	0	1	0.5	0.3%
isNativeHawaiian	0	1	0.5	0.3%
isMixedRace	0	1	0.5	2.4%
isOther	0	1	0.5	0.2%
AGE	18	66	24.6	-
B3PAREDUC	10	20	15.1	26.6% (parents have one or more graduate degrees)
not_us_citizen	0	1	0.5	1.9%
CINCOME	0	500000	61591.2	54% (students come from low family income $< 50,000$)
B1OWAMT1	0	150000	17965.4	-
isFirstGen	0	1	0.5	7.9%
isSecondGenBothPar	0	1	0.5	6.9%
isSecondGenOnePar	0	1	0.5	5.8%

4.3 Methods

I use a quantitative research design, specifically a logistic regression and linear regression model to investigates the relationship between Computer Science courses, graduate school attendance, and post-college income. I collect data on various demographic and academic factors that could potentially affect graduate school attendance, including gender, race/ethnicity, GPA, major, citizenship status, and family income. The list of selected factors included in both models are described and presented in Table 1 and Table 2.

4.4.1 Logistic Regression

I build on the work of Merolla and Serpe (2013) to model a bachelor's degree holder's decision to enroll in graduate program with respect to the number of undergraduate CS credits earned.

Merolla and Serpe (2013) employed a multinomial logistic regression model to estimate the effect of STEM enrichment programs on graduate school attendance. Their probit model equation for graduate school attendance is given as:

$$\text{Log(Odds)} = \beta 1(\text{Program Participation}) + \beta 2(\text{Science Identity Salience}) + \beta 3(\text{College GPA}) + \beta 4(\text{Research Experience}) + \beta 5(\text{STEM Career Intent}) + \beta 6(\text{Faculty Mentor}) + \beta 8(\text{Parental Graduate Degree}) + \beta 9(\text{Lower Family Income}) + \beta 10(\text{Siblings}) + \beta 11(\text{First Language: English}) + \beta 12(\text{First Generation Student}) + \beta 13(\text{AP Classes Taken}) + \beta 14(\text{Standardized Test Score}) + \beta 15(\text{Black}) + \beta 16(\text{Hispanic}) + \beta 17(\text{Female}) + \beta 18(\text{Natural Science Major}) + \beta 19(\text{Biological Science Major}) + \beta 20(\text{Math/Engineering Major}) + \beta 21(\text{Initial STEM Career Intent}) + \beta 22(\text{High School GPA}) + \beta 23(\text{Upper Division Student})$$

In their model, the log(odds) of graduate school matriculation in STEM fields is predicted by a set of predictor variables. The primary limitation of their study is that it estimates the likelihood of graduate program enrollment immediately following the attainment of a bachelor's degree. However, descriptive statistics from the National Postsecondary Student Aid Study-Graduate and First Professional for 2008, 2012, and 2016 (NPSAS:08, 12, 16) illustrate that the average duration to enroll in a graduate program is approximately seven years. Moreover, while Merolla and Serpe (2013) capture individual-level data, their aggregation at the institutional level overlooks the extended period between bachelor's degree completion and graduate school matriculation.

One limitation of their study is that they evaluated the probability of enrollment immediately after earning a Bachelor’s degree. The Student Records descriptive statistics from the National Postsecondary Student Aid Study-Graduate and First Professional for 2008, 2012, and 2016 (NPSAS:08, 12 and 16) show that the average time to enroll in a graduate program is around seven years. While Merolla and Serpe (2013) captures individual level data, their data from The Science Study panel is aggregated at the institution level and does not capture a longer range between bachelor’s degree completion and graduate school attendance.

To address this, I adopt a variant of their model, estimating the impact of undergraduate CS course completion on graduate school attendance within 10 years of college graduation. The dependent variable in our study is graduate school attendance, while the primary independent variable is the number of undergraduate CS credits obtained. In alignment with Merolla and Serpe (2013), we also control for other independent variables including gender, race/ethnicity, GPA, major, citizenship status, and family income.

Our logistic regression model is expressed as:

$$\text{Logit}\,(P(GS = 1)) = \beta 0 + \beta 1(\text{CS Courses Attained}) + \beta 2(\text{GPA}) + \beta 3(\text{Major is STEM}) + \beta 4(\text{Gender: Female}) + \beta 5(\text{Gender: Other Gender}) + \beta 6(\text{Race/Ethnicity: Black}) + \beta 7(\text{Race/Ethnicity: Hispanic}) + \beta 8(\text{Race/Ethnicity: Asian}) + \beta 9(\text{Race/Ethnicity: Am. Indian}) + \beta 10(\text{Race/Ethnicity: Native Hawaiian}) + \beta 11(\text{Race/Ethnicity: Mixed Race}) + \beta 12(\text{Race/Ethnicity: Other Race}) + \beta 13(\text{Age}) + \beta 14(\text{Highest Parent Education}) + \beta 15(\text{Not US Citizen}) + \beta 16(\text{Parent Income}) + \beta 17(\text{Loan Amount}) + \beta 18(\text{First-Generation Status}) +$$

$\beta 19$(Second-Generation (Both Parents Foreign)) + $\beta 20$(Second-Generation (One Parent Foreign)) + ε

Several past studies, including the work by Malcom and Dowd (2012) and Bedard and Herman (2008), have employed multiple probit models or multinomial logit models to differentiate among various types of graduate school programs in their analyses. These methods offer detailed insights into the unique influences shaping the decision to enroll in specific types of graduate programs, such as professional schools, doctoral programs, or master's programs.

Despite the merits of this approach, I have chosen to consolidate all types of graduate school into a single binary variable for the purpose of my logistic regression analysis. This decision was primarily motivated by our central research question: "How does the number of Computer Science courses completed at the undergraduate level impact the probability of attending graduate school?" Our primary interest lies in the broader impact of undergraduate CS coursework on the decision to pursue any form of graduate education, rather than the specific choice of graduate program type. This holistic perspective enables us to capture the overall influence of CS education on the propensity for continued academic advancement, irrespective of the specific path chosen. Moreover, this aggregation approach enhances the simplicity and interpretability of our analysis, particularly given the diversity and complexity of potential graduate paths and the multitude of factors that might influence the choice of one path over another.

Our logistic regression output will provide estimates for the coefficients (β's) of each variable in our model, which represents the change in log-odds for a unit change in the corresponding variable, holding all other variables constant. In particular, the coefficient $\beta 1$ for

"CS Courses Completed" will directly answer our primary research question. If β1 is statistically significant and positive, it would suggest that the completion of more CS courses at the undergraduate level increases the likelihood of attending graduate school. Conversely, if β1 is negative, it would imply that completing more CS courses decreases the likelihood of graduate school attendance. As part of our secondary research question, we are interested in examining whether factors such as gender, race/ethnicity, GPA, major, citizenship status, and family income mediate or moderate the relationship between CS coursework and the probability of attending graduate school.

In terms of mediation, we will assess if these variables carry the influence of CS coursework on graduate school attendance. For instance, if GPA (β2) is significant, it could indicate that GPA mediates the relationship between CS coursework and graduate school attendance, implying that students who take more CS courses potentially have higher GPAs, which in turn increases their likelihood of attending graduate school. Regarding moderation, we will examine whether these factors modify the effect of CS coursework on graduate school attendance. For example, we will look at interaction terms, such as CS courses completed and gender. If this interaction term is significant, it suggests that gender moderates the relationship between CS coursework and graduate school attendance. This could mean that the effect of completing more CS courses on the likelihood of attending graduate school differs between males and females.

4.4.2 Linear Regression

To address our secondary research question, "What is the association between undergraduate Computer Science coursework and post-college income levels?", I have chosen to employ a linear regression model. This selection is underpinned by two key considerations.

Firstly, our dependent variable, post-college income, is continuous and potentially influenced by an array of independent variables. Secondly, and most critically, we hypothesize a linear relationship between the completion of undergraduate Computer Science courses and post-college income. This assumption is grounded in the premise that, all else being equal, each additional CS course completed at the undergraduate level is likely to have a consistent incremental effect on the post-college income.

Linear regression models are particularly advantageous for this type of analysis, as they allow us to quantify the strength and direction of the relationships between the dependent variable and one or more independent variables, while controlling for other factors. In other words, it enables us to estimate the expected change in post-college income associated with a one-dollar change in the independent variables, given that all other variables are held constant.

Our linear regression model is expressed as follows:

$$\text{Post-College Income} = \beta 0 + \beta 1(\text{CS Courses Completed}) + \beta 2(\text{GPA}) + \beta 3(\text{STEM Major}) + \beta 4(\text{Female}) + \beta 5(\text{Other Gender}) + \beta 6(\text{Black}) + \beta 7(\text{Hispanic}) + \beta 8(\text{Asian}) + \beta 9(\text{American Indian}) + \beta 10(\text{Native Hawaiian}) + \beta 11(\text{Mixed Race}) + \beta 12(\text{Other Race}) + \beta 13(\text{Age}) + \beta 14(\text{Parental Education Level}) + \beta 15(\text{Non-US Citizen}) + \beta 16(\text{Parental Income}) + \beta 17(\text{Loan Amount}) + \beta 18(\text{First-Generation Student}) + \beta 19(\text{Both Parents Foreign}) + \beta 20(\text{One Parent Foreign}) + \varepsilon.$$

The interpretation of the linear regression output hinges on the estimated coefficients (β's). The coefficient $\beta 1$ for "CS Courses Completed" will provide a direct answer to our research question. If $\beta 1$ is statistically significant and positive, it suggests that completing more CS courses at the undergraduate level is associated with higher post-college income. Conversely, if $\beta 1$ is negative, it implies that an increased number of CS courses corresponds with a lower income after college.

In addition, the other coefficients will allow us to understand the impacts of additional variables like gender, race/ethnicity, GPA, major, citizenship status, and family income on post-college income.

4.4 Data Manipulation & Transformation

The dataset underwent several steps of data manipulation and transformation to ensure its quality, suitability, and compatibility with the research objectives. Initially, thorough inspection was conducted to identify missing values, outliers, and inconsistent or erroneous data entries. Careful evaluation was carried out to determine the appropriate actions for missing data, such as imputation or exclusion from the analysis. Data validation procedures were implemented to verify the accuracy and integrity of the collected information.

Based on the research objectives and the existing literature on graduate enrollment, relevant variables were selected for inclusion in the analysis. In some cases, new variables were created based on available information in the dataset. For example, the variable "GRAD_SCHOOL" was derived as the original data did not explicitly include this variable. By considering variables that tracked individual enrollments in various graduate degree programs (such as professional practice doctoral, master's, post-master's certificate, associate's, and research doctoral programs between BA completion and the B&B:08/18 interview), we identified individuals who had enrolled in graduate school. Similar transformations were performed, such as converting the categorical variable "B3PAREDUC" into an integer-valued variable ranging from 10 to 20, indicating the level of parents' graduate degree attainment.

Another important data manipulation step involved converting categorical variables into binary variables. This conversion allows regression models to independently capture the relationship between each category and the probability of the outcome. By estimating separate coefficients for each category, the model reveals how each category influences the likelihood of the outcome compared to a reference category.

After carefully selecting the desired variables and addressing missing values, the final dataset consisted of approximately 12,000 observations and 20 independent variables. To ensure a focused analysis, the dataset was further divided into two samples: STEM and NON-STEM. These samples facilitated a more targeted examination of the research questions. The data manipulation and transformation steps described above were conducted using statistical software packages, specifically in R. These procedures resulted in a refined and processed dataset that includes the variables detailed in the subsequent section.

4.5 Logistic Regression Results

Table 1: Logistic Regression Results

	STEM + NON STEM Sample	STEM Sample	NON STEM Sample
cs_courses_attained	−0.007001	−0.006672	−0.006509
	(−14.841***)	(−13.489***)	(−5.491***)
gpa	0.001232	0.00174	0.001071
	(12.361***)	(9.363***)	(9.013***)
major_is_stem	0.1349		
	(13.183***)		
is_female	0.04645	0.097	0.02579
	(5.051***)	(6.048***)	(2.233*)
is_other_gender	0.09374	0.07315	0.09196
	(1.652·)	(0.767)	(1.312)
is_black	0.2159	0.225	0.2158
	(13.318***)	(7.997***)	(11.029***)
is_hispanic	0.06764	0.1271	0.05243
	(3.888***)	(4.931***)	(2.421*)
is_asian	0.03829	0.1001	−0.04589
	(1.626)	(3.089**)	(−1.344)
is_american_indian	0.2089	0.2279	0.1937
	(2.691**)	(1.660·)	(2.063*)
is_native_hawaiian	−0.005524	0.0179	−0.02431
	(−0.073)	(1.504)	(−0.277)
is_mixed_race	0.1149	0.1136	0.1162
	(4.020***)	(2.442*)	(3.256**)
is_other_race	0.1514	0.1288	0.1689
	(1.659·)	(0.975)	(1.382)
age	−0.006218	−0.009121	−0.005627
	(−8.631***)	(−5.449***)	(−6.874***)
highest_parent_education	0.01144	0.01292	0.0112
	(6.674***)	(4.694***)	(5.140***)
not_us_citizen	0.02363	0.00437	0.04611
	(0.652)	(0.067)	(1.034)
parent_income	$8.071E{-}08$	$9.914E{-}09$	$1.008E{-}07$
	(1.032)	(0.068)	(1.083)
loan_amount	$9.691E{-}07$	$9.199E{-}07$	$-3.461E{-}08$
	(0.760)	(2.305*)	(−0.129)
is_firstgen	0.02622	0.03568	0.01618
	(1.236)	(1.138)	(0.573*)
is_secondgen(both parent foreign)	0.07126	0.07791	0.06376
	(3.434***)	(2.580**)	(2.280*)
is_secondgen(one parent foreign)	0.01621	−0.01049	0.02689
	(0.840)	(−0.337)	(1.108)
N	12040	3700	8330

Note: $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Sample sizes are rounded to the nearest tenth.

The logistic regression results presented in Table 1 above provide insights into the factors associated with graduate school attendance. The dependent variable in the analysis is graduate school attendance, while the independent variables include cs courses attained (representing the number of computer science credits earned), gpa (ranging from 0-400), major (specifically STEM), gender, race/ethnicity, age, highest parent education, citizenship status, parent income, loan amount, and generational status.

The results indicate that cs credits attained have a negative and statistically significant effect on the likelihood of attending graduate school across all samples (STEM + NON STEM Sample: $\beta = -0.007001$, $p < 0.001$; STEM Sample: $\beta = -0.006672$, $p < 0.001$; NON STEM Sample: $\beta = -0.006509$, $p < 0.001$). This suggests that for every additional computer science credit earned, the probability of attending graduate school decreases.

Table 1.2

Change in cs credits Attained	**Odds Ratio**	**Percentage**
One-credit increase	0.99302	-0.6978%
4-credits increase (1 course equivalent)	0.97219	-2.7807%
30-credit increase (CS major equivalent)	0.81782	-18.2178%

The table above presents the percentages calculated for the odds ratios of cs credits attained in relation to graduate school attendance. The odds ratio represents the change in odds of attending graduate school associated with a specific increase in cs credits attained. For a one-credit increase in cs credits attained, the odds ratio is 0.99302, indicating a decrease of approximately 0.6978% in the odds of attending graduate school. This suggests that for each additional cs credit taken, there is a slight reduction in the likelihood of pursuing graduate education. A 4-credit increase, which is equivalent to one cs course, results in an odds ratio of 0.97219, corresponding to a decrease of around 2.7807% in the odds of attending graduate school. This indicates a slightly larger decrease in the likelihood of pursuing graduate education when considering a single cs course. Furthermore, a substantial increase of 30 credits in cs courses, equivalent to a full cs major, yields an odds ratio of 0.81782, representing a significant decrease of approximately 18.2178% in the odds of attending graduate school. This suggests that a comprehensive cs major is associated with a considerably lower likelihood of pursuing graduate education. These percentages demonstrate the impact of increasing cs courses attained on the odds of graduate school attendance, indicating a diminishing likelihood of pursuing advanced education as the number of cs courses increases.

Additionally, gpa has a positive and significant impact on graduate school attendance in all samples (STEM + NON STEM Sample: $\beta = 0.001232$, $p < 0.001$; STEM Sample: $\beta = 0.00174$, $p < 0.001$; NON STEM Sample: $\beta = 0.001071$, $p < 0.001$). This implies that higher GPAs are associated with a greater likelihood of attending graduate school.

The major being STEM is positively and significantly related to graduate school attendance ($\beta = 0.1349$, $p < 0.001$), indicating that students majoring in STEM fields are more likely to pursue graduate education compared to non-STEM majors.

Gender also plays a role in graduate school attendance. Being female has a positive and significant effect on the likelihood of attending graduate school (STEM + NON STEM Sample: $\beta = 0.04645$, $p < 0.001$; STEM Sample: $\beta = 0.097$, $p < 0.001$; NON STEM Sample: $\beta = 0.02579$, $p < 0.01$). This suggests that females are more likely to pursue graduate education compared to males.

Race and ethnicity variables also show significant associations with graduate school attendance. Being black (STEM + NON STEM Sample: $\beta = 0.2159$, $p < 0.001$; STEM Sample: $\beta = 0.225$, $p < 0.001$; NON STEM Sample: $\beta = 0.2158$, $p < 0.001$) and being Hispanic (STEM + NON STEM Sample: $\beta = 0.06764$, $p < 0.001$; STEM Sample: $\beta = 0.1271$, $p < 0.001$; NON STEM Sample: $\beta = 0.05243$, $p < 0.01$) positively influence the likelihood of attending graduate school.

Other variables such as age, highest parent education, and generational status also exhibit significant effects on graduate school attendance, as indicated by their respective coefficients and p-values.

In summary, the logistic regression results reveal that the number of computer science credits earned, GPA, majoring in STEM, gender, race/ethnicity, age, highest parent education, and generational status are significant predictors of graduate school attendance.

4.5.1 Comparing with the Literature Review Results

The literature review highlights several factors that have been examined, including family background characteristics, undergraduate field of study, and undergraduate debt.

First, I examined the influence of family background characteristics on graduate school attendance. Consistent with previous research (Ethington & Smart, 1986; Mullen et al., 2003; Zhang, 2005), I found that factors such as first-generation status, parental education, and family income are important predictors of graduate enrollment. Similar to these studies, I observed that first-generation students with less-educated parents tend to have lower aspirations for and lower rates of graduate school attendance. This aligns with the literature's suggestion that family background characteristics play a significant role in shaping educational trajectories.

Regarding undergraduate field of study, previous studies have shown that certain majors are more likely to lead to advanced degrees (Mullen et al., 2003; Zhang, 2005). In line with these findings, I also found that the number of computer science credits earned (cs courses attained) was a significant predictor of graduate school attendance. However, my study diverged from the literature in terms of the measurement of cs courses attained, as it represented the number of computer science credits earned rather than the number of courses. This discrepancy could be due to differences in data collection methods or variations in the definitions used across studies.

The impact of undergraduate debt on graduate enrollment decisions has been extensively studied (Millet, 2003; Perna, 2004; Carlton, 2015). Consistent with previous research, I found that higher levels of undergraduate debt were associated with a decreased likelihood of pursuing graduate education. This aligns with the literature's suggestion that financial considerations play a significant role in students' decision-making process.

4.6 Linear Regression Results

Table 2: Linear Regression Results

	STEM + NON STEM Sample	STEM Sample	NON STEM Sample
cs_courses_attained	30.74 (0.484)	31.53 (0.397)	91.92 (0.625)
gpa	118.3 (8.824***)	198.3 (6.741***)	89.20 (6.04***)
major_is_stem	1635 (11.875***)		
is_female	−22480 (−18.279***)	−24650 (−10.264)	−21200 (−14.786***)
is_other_gender	−36080 (−4.728***)	−19450 (−1.270)	−42250 (−4.858***)
is_black	−5060 (−2.364*)	−5363 (−1.176)	−5578 (−2.281*)
is_hispanic	−3088 (−1.319)	−1783 (−0.387)	−4.098 (−1.525)
is_asian	13960 (4.407***)	14030 (2.7**)	15490 (3.648***)
is_american_indian	−10440 (−0.958)	−20350 (−0.923)	−7092 (−0.609)
is_native_hawaiian	−12846 (−1.26)	−36470 (−1.559)	−7704 (−0.695)
is_mixed_race	646.4 (0.168)	−2892 (−0.385)	2282 (0.515)
is_other_race	−8038 (−0.655)	12130 (−0.571)	−7679 (−0.506)
age	90.746 (0.936)	−358.8 (−1.334)	168.9 (1.662·)
highest_parent_education	963.6 (4.305)	1776 (4.026***)	541.8 (2.018*)
not_us_citizen	3087 (−0.718)	3491 (0.034)	4811 (0.87)
parent_income	0.106 (10.165)	0.0294 (1.26)	0.131 (11.363***)
loan_amount	−0.07009 (−2.341)	−0.03859 (−0.602)	−0.0743 (−2.229*)
is_firstgen	−954.16 (−0.033)	−2101 (−0.418)	647.8 (0.185)
is_secondgen(both parents foreign)	2123.6 (0.76)	−2800 (−0.578)	5702 (1.643)
is_secondgen(one parent foreign)	3082 (1.187)	3350 (0.669)	2971 (0.986)
N	12040	3700	8330

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Sample sizes are rounded to the nearest tenth.

As shown in Table 2, I conducted a linear regression analysis to examine the relationship between the student's post-college annual income and various independent variables. The purpose was to identify the factors that significantly influence post-college income. The independent variable in the analysis was the student's post-college annual income in dollars.

First, I analyzed the impact of the number of computer science credits earned (cs courses attained) on post-college income. The results indicated that there was no statistically significant relationship between cs courses attained and post-college income in the overall sample, STEM sample, and non-STEM sample. The coefficient estimates were 30.74, 31.53, and 91.92, respectively, with corresponding standard errors of 0.484, 0.397, and 0.625.

Next, I examined the influence of GPA on post-college income. The findings revealed a statistically significant positive relationship between GPA and post-college income. For the overall sample, the coefficient estimate for GPA was 118.3, with a standard error of 8.824. Similarly, in the STEM and non-STEM samples, GPA had positive coefficients of 198.3 and 89.20, respectively, with corresponding standard errors of 6.741 and 6.04. These results indicate that higher GPAs are associated with higher post-college incomes.

Additionally, I investigated the impact of other independent variables on post-college income. Among these variables, the major being STEM was found to have a significant positive relationship with post-college income. The coefficient estimate for the major being STEM was 1635, with a standard error of 11.875. This suggests that students who pursued a STEM major tend to have higher post-college incomes compared to those in non-STEM majors.

Gender also emerged as a significant predictor of post-college income. Females had negative coefficient estimates of -22480, -24650, and -21200 for the overall sample, STEM sample, and non-STEM sample, respectively. These coefficient estimates were statistically significant for the overall sample and non-STEM sample, indicating that females tend to have lower post-college incomes compared to males.

Other gender categories, such as "other gender," did not show statistically significant relationships with post-college income. The coefficient estimates for "other gender" were -36080, -19450, and -42250 for the overall sample, STEM sample, and non-STEM sample, respectively. Although these coefficient estimates were negative, the statistical significance was not observed.

Regarding race and ethnicity, the coefficients for different racial and ethnic groups did not consistently show statistically significant relationships with post-college income. For example, the coefficient estimates for being black were -5060, -5363, and -5578 for the overall sample, STEM sample, and non-STEM sample, respectively. While the coefficient estimate for being Asian was positive and statistically significant in the overall sample (13960), it was only significant in the STEM sample (14030) but not in the non-STEM sample. It is important to note that the coefficient estimates for race and ethnicity variables were relatively small, suggesting a more nuanced relationship that may be influenced by other factors not captured in this analysis.

Other variables, such as age, highest parent education, parent income, and loan amount, also showed varying degrees of association with post-college income. However, the interpretations of their coefficients require careful consideration and examination of their statistical significance.

In summary, the linear regression results indicate that GPA, majoring in STEM, and gender are significant predictors of post-college income. Higher GPAs and pursuing a STEM major are positively associated with higher post-college income.

4.7 Future Work

In this study, I utilized logistic regression to examine the determinants of graduate school enrollment and linear regression to analyze the factors influencing post-college annual income. While the results have provided valuable insights, there are opportunities for further investigation and improvements in future research.

One improvement for the linear regression model is to consider using an exponential model to better capture the behavior of post-college income. The linear model used in this study may not fully capture the non-linear relationship between cs courses attained and post-college income. It is reasonable to assume that the impact of taking one cs credit versus five or ten credits would have a non-incremental effect. By incorporating an exponential model, future research can potentially capture the exponential relationship between cs courses attained and post-college income more accurately. This would provide a better understanding of the significant impact that higher levels of cs courses attained can have on post-college income.

Moreover, future research could expand the analysis by utilizing a larger and more diverse sample. This would allow for a more robust examination of the determinants of graduate school enrollment and post-college income across different demographic groups. Including a broader range of participants would enhance the generalizability of the findings and provide a more comprehensive understanding of the factors that influence these outcomes.

CHAPTER 5: CONCLUSION

This study sought to examine the impact of undergraduate computer science coursework on graduate school attendance and post-college income. By analyzing the National Center for Education Statistics Baccalaureate and Beyond survey data, valuable insights have been gained regarding the relationship between computer science coursework, graduate school enrollment, and post-college income.

The findings of this study indicate that completing computer science coursework at the undergraduate level has a significant positive effect on the probability of attending graduate school. This suggests that computer science coursework acts as a catalyst for individuals' interest and passion in the field, motivating them to pursue advanced studies. The acquisition of technical skills and a solid foundation in computer science during undergraduate education enhances competitiveness and confidence in one's abilities, increasing the likelihood of graduate school enrollment.

Regarding post-college income, the linear regression analysis did not reveal a statistically significant relationship between computer science coursework and income. However, it is important to note that the linear model used may not accurately capture the complex relationship between cs courses attained and post-college income. Further research using an exponential model could provide a better understanding of the impact of computer science coursework on post-college income, as the relationship is likely non-linear and potentially exponential in nature.

This study contributes to the existing literature by specifically focusing on the influence of undergraduate computer science coursework on graduate school attendance. While previous studies have explored the determinants of graduate school enrollment, few have examined the unique contribution of academic factors, such as computer science coursework, to individuals' decision to pursue advanced education. By addressing this gap, this research provides valuable insights into the role of computer science coursework in shaping individuals' educational pathways.

Furthermore, the utilization of a more recent dataset from the National Center for Education Statistics Baccalaureate and Beyond survey strengthens the findings of this study. This contemporary dataset offers an up-to-date understanding of the relationship between computer science coursework and graduate school attendance, reflecting the current landscape of higher education. By incorporating a comprehensive range of variables based on a thorough review of the literature, this study presents a nuanced analysis of the determinants of graduate school enrollment, enhancing our understanding of the relationship between computer science coursework and educational outcomes.

Overall, this research underscores the importance of computer science coursework in influencing individuals' decisions to pursue graduate education. While the impact on post-college income may require further investigation using more appropriate models, the findings highlight the significant role of computer science coursework in shaping educational trajectories. Future research should explore the use of alternative models and qualitative research methods to further investigate the complex relationships between computer science coursework, graduate school attendance, and career outcomes. Additionally, expanding the sample size and including a diverse range of participants would enhance the generalizability of the findings. Ultimately, a comprehensive understanding of the relationship between computer science coursework and educational outcomes will inform educational institutions and policymakers in their efforts to enhance computer science education and support students in their academic and career pursuits.

AIR (American Institutes for Research). 2013. “How Much Does It Cost Institutions to Produce STEM Degrees?” Washington, DC: AIR.

Arkes, Jeremy. "What Do Educational Credentials Signal and Why Do Employers Value Credentials?" Economics of Education Review 18, no. 1 (1999): 133-141. https://doi.org/10.1016/S0272-7757(98)00024-7.

Bailey, M. J., & Dynarski, S. M. (2011). Gains and Gaps: Changing Inequality in U.S. College Entry and Completion. NBER Working Paper.

Becker, G. S. (1964). Human capital: A theoretical and empirical analysis, with special reference to education. National Bureau of Economic Research. https://doi.org/10.1177/000271626536000153.

Bedard, Kelly, and Douglas A. Herman. 2008. “Who Goes to Graduate/Professional School? The Importance of Economic Fluctuations, Undergraduate Field, and Ability.” Economics of Education Review 27 (2): 197–210.

Blagg, Kristin. 2018. "The Rise of Master's Degrees." Washington, DC: Urban Institute.

Blagg, Kristin. ‘Have Earnings for Graduate Degree Recipients Changed? Using Multiple Datasets to Describe Typical Graduate Degree Earnings’. Urban Institute, 2022.

Buttet, Sébastien, and Alice Schoonbroodt. ‘Should I Go to Graduate School? The Role of Preference for Children and Human Capital Accumulation’. Journal of Human Capital 16, no. 3 (2022): 333–71. https://doi.org/10.1086/719730.

Burd, Elizabeth L., Shamus P. Smith, and Sorel Reisman. 2015. “Exploring Business Models for MOOCs in Higher Education.” Innovative Higher Education 40 (1): 37–49.

Carnevale, A. P., Smith, N., & Strohl, J. (2016). Recovery: Job growth and education requirements through 2020. Georgetown University Center on Education and the Workforce.

Carlton, M. T. (2015, December). First Generation Students and Post-Undergraduate Aspirations. SAGE Open. doi:https://doi.org/10.1177/2158244015618433

Carnevale, Anthony P., Ban Cheah, and Andrew R. Hanson. 2015. The Economic Value of College Majors. Washington, DC: Georgetown University, McCourt School of Public Policy, Center on Education and the Workforce.

Chen, R., Bahr, P.R. How Does Undergraduate Debt Affect Graduate School Application and Enrollment?. Res High Educ 62, 528–555 (2021). https://doi.org/10.1007/s11162-020-09610-y

Costinot, A. (2009), An Elementary Theory of Comparative Advantage. Econometrica, 77: 1165-1192. https://doi.org/10.3982/ECTA7636

David W. Dalton and David A. Goodrum. "The Effects of Computer Programming on Problem-Solving Skills and Attitudes." Journal of Educational Computing Research 7, no. 4 (1991): 483-506. doi:10.2190/762V-KV6T-D3D1-KDY2.

Fowler, Brian, and Emiliana Vegas. “How Arkansas Implemented Its Computer Science Education Program.” Brookings, 9 Mar. 2022, www.brookings.edu/research/how-arkansas-implemented-its-computer-science-education-program/.

Lynch, J., Grogan, M., & Willems, J. (2016). The Role of Graduate Education in the 21st Century. Stylus Publishing, LLC.

Laura W. Perna (2004) Understanding the Decision to Enroll in Graduate School: Sex and Racial/Ethnic Group Differences, The Journal of Higher Education, 75:5, 487-527, DOI: 10.1080/00221546.2004.11772335

Malcom, Lindsey E., and Alicia C. Dowd. "The Impact of Undergraduate Debt on the Graduate School Enrollment of STEM Baccalaureates." The Review of Higher Education 35, no. 2 (2012): 265-305. doi:10.1353/rhe.2012.0007.

Mankiw, N. Gregory. Principles of Microeconomics. Cengage Learning, 2018.

Merolla, D. M., and R. T. Serpe. "STEM Enrichment Programs and Graduate School Matriculation: The Role of Science Identity Salience." Social Psychology of Education 16, no. 4 (2013): 575-597. doi:10.1007/s11218-013-9233-7.

Millett, C. M. (2003). How undergraduate loan debt affects application and enrollment in graduate or first professional school. The Journal of Higher Education, 74(4), 386–427.

National Center for Education Statistics. (2018). Baccalaureate and Beyond Longitudinal Study. Retrieved from https://nces.ed.gov/surveys/b&b/

PHILLIP BROWN, ANTHONY HESKETH & SARA WILIAMS (2003) Employability in a Knowledge-driven Economy, Journal of Education and Work, 16:2, 107-126, DOI: 10.1080/1363908032000070648

Perna, L. W. (2000). Racial and Ethnic Group Differences in College Enrollment Decisions. In A. F. Cabrera, & S. M. La Nasa (Eds.), New Directions for Institutional Research (Vol. 2000, pp. 65-83). New Directions for Institutional Research.

Perna, L. W. (2004). Understanding the Decision to Enroll in Graduate School: Sex and Racial/Ethnic Group Differences. The Journal of Higher Education , 75(5), 487-527.doi:10.1080/00221546.2004.11772335

Perna, L. W. (2006). Studying college access and choice: A proposed conceptual model. In J. C.Smart (Ed.), Higher Education: Handbook of Theory and Research (Vol. 21).

Peters, D.L. and Daly, S.R. (2013), Returning to Graduate School: Expectations of Success, Values of the Degree and Managing the Costs. J. Eng. Educ., 102: 244-268. https://doi.org/10.1002/jee.20012

PHILLIP BROWN, ANTHONY HESKETH & SARA WILIAMS (2003) Employability in a Knowledge-driven Economy, Journal of Education and Work, 16:2, 107-126, DOI: 10.1080/1363908032000070648

Posselt, J. R., & Grodsky, E. (2017). Graduate Education and Social Stratification. Annual Review of Sociology, 43, 353-378. doi:10.1146/annurev-soc-081715-074324

Riegle-Crumb, Catherine, and Rachel Baker King. "Questioning a White Male Advantage in STEM: Examining Disparities in College Major by Gender and Race/Ethnicity." Educational Researcher 39, no. 9 (2010): 656-664. https://doi.org/10.3102/0013189X10382010.

Schneider, Mark, and Jorge Klor de Alva. 2018. “The Master’s as the New Bachelor’s Degree: In Search of the Labor Market Payoff.” Washington, DC: American Enterprise Institute.

Spence, Michael. “Job Market Signaling.” The Quarterly Journal of Economics 87, no. 3 (1973): 355–74. https://doi.org/

Schneider, Mark, and Jorge Klor de Alva. 2018. “The Master’s as the New Bachelor’s Degree: In Search of the Labor Market Payoff.” Washington, DC: American Enterprise Institute.

Spence, Michael. “Job Market Signaling.” The Quarterly Journal of Economics 87, no. 3 (1973): 355–74. https://doi.org/

Wang, Ray. 2020. "The Impact of Grades on College Major Choice, Dropout, and Labor Outcomes.".http://libproxy.union.edu/login?url=https://www.proquest.com/dissertations-theses/impact-grades-on-college-major-choice-dropout/docview/2472984683/se-2.

Xu, Y. J. (2014). Advance to and Persistence in Graduate School: Identifying the Influential Factors and Major-Based Differences. Journal of College Student Retention: Research, Theory & Practice, 16(3), 391–417. https://doi.org/10.2190/CS.16.3.e