UNDERSTANDING THE PREDICTORS OF STUDENT LOAN DEFAULT FOR
COMMUTER STUDENTS AT A NONRESIDENTIAL MASTER'S COMPREHENSIVE
CAMPUS

By

Daniel Z. Merian

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Higher, Adult, and Lifelong Education—Doctor of Philosophy

2021

ABSTRACT

UNDERSTANDING THE PREDICTORS OF STUDENT LOAN DEFAULT FOR COMMUTER STUDENTS AT A NONRESIDENTIAL MASTER'S COMPREHENSIVE CAMPUS

By

Daniel Z. Merian

In the 21st century, more students enroll in higher education and take federal loans to defer the cost of attendance resulting in average levels of borrowing steadily increasing. In the same timeframe, there is an increase in the number of students entering repayment for their federal loans and an increase in the proportion of individuals defaulting on their repayments. Individual institutions are responsible for the students who borrow money to attend their institution. As such, individual campuses are interested in knowing which specific factors affect their students' ability to repay their loans. Therefore, via a quantitative case study, I sought to understand if local, institution-specific data could improve the timing of the estimates of default probabilities at a nonresidential campus. My case study allowed me to examine how institutional-level data could inform institutional decision-making by showing which pre-college student characteristics, college experience by semester enrolled variables, and post-attendance factors associated with default for my subject institution. Specifically, by incorporating these institutional-level variables in an organized way (i.e., by type of activity or by semester), I could determine how early in a student's career I could identify them as "at-risk" of default. My findings show that institution-specific, student-level data at my subject institution provides a template to predict default sooner than the traditional default measures and provides a framework for other institutions to apply to their student populations. My results suggest existing default

research may be missing an important set of variables in the models, variables measuring

students' college experiences.

For Laurie, Kaiden, Ezra, and Declan.

ACKNOWLEDGMENTS

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

Enrolling in higher education is one of the strongest investments individuals can make for themselves. According to a report published by the United States President's Council of Economic Advisors, the median worker with a bachelor's degree, throughout a career, earns nearly $1 million more than a similar worker with just a high school diploma (Council of Economic Advisors, 2016). Further, bachelor degree recipients experience lower levels of unemployment and have increased odds of moving up the economic ladder (Council of Economic Advisors, 2016). Overall, data suggest that enrolling in higher education can result in positive economic outcomes for students.

Although the literature suggests higher education is a sound investment, undergraduate students who enroll in higher education, on average, cannot afford the entire cost of attendance (Council of Economic Advisors, 2016). As such, students need financial aid to help reduce and/or defer the cost of their education to a later date. One mechanism students often use to defer the cost of higher education is student loans, the majority of which are federal loans (Baum et al., 2017). In fact, in 2020–21 alone, over 6.8 million undergraduate students utilized federal loans to defer the cost of attendance. Further, the number of borrowers and levels of borrowing of federal loans has reached record levels (Ma & Pender, 2021). In the third quarter of 2021, there were 42.9 million total borrowers with a combined outstanding balance of approximately $1.6 trillion (Federal Reserve Bank of New York, 2021). This is a 17% increase in the total number of borrowers from 2011 and a 51% increase from 2007 (Compiled from Department of Education, 2021, and author's calculations). The amount of money borrowed has grown at an even higher rate than the number of borrowers. In 2007, total outstanding debt was $547 billion, but in 2003, the outstanding balance was just $253 billion (Federal Reserve, 2021). In 2021, the current total

loan debt is 2.9 times more than 2007 and approximately 6.3 times more than 2003 (Compiled from Department of Education, 2021, and author's calculations).

More individuals borrowing has resulted in more individuals entering loan repayment. As both average loan amounts and the number of students entering repayment have increased, so has the number of borrowers' defaulting (Baum et al., 2017; Council of Economic Advisors, 2016). Just because an individual enters repayment does not mean he or she will default on the loan. However, the potential problem is as the number of borrowers and levels of outstanding debt increases, so does the risk of loan default. As demonstrated by Federal Reserve Bank data, in 2021, 5.5 million borrowers were in default compared to 1.9 and 1.1 million borrowers in 2008 and 2003, respectively (Federal Reserve Bank of New York, 2021). This represents default rates of 6% and 7% in 2003 and 2008, respectively, compared to approximately 11% of borrowers defaulting in 2021 (Federal Reserve Bank of New York, 2021). This is an issue for a multitude of reasons. For the defaulters, debt is not forgiven through bankruptcy, it adversely affects their credit scores, and adversely affects their ability to create wealth through such means as purchasing a home. Society is adversely affected by the defaulting of borrowers because the outstanding balances owed back to the Federal government are not being repaid.

To better understand the loan default phenomenon, researchers have studied the predictors of defaulting. Findings suggest strong predictors of default include whether a student graduates from higher education and the borrower's background characteristics (Grosset al., 2009). For instance, non-completers are more likely to default than their graduated peers (Baum et al., 2017). Parental education is a significant indicator of student default (Volkwein et al., 1998; Volkwein & Szelest, 1995), and students of color are more likely to default than White students (Volkwein & Szelest, 1995; Woo, 2002). With the knowledge that specific background

characteristics are strong predictors of default, the examination of unstudied characteristics is warranted.

While national studies have identified the predictors of default, are the findings from a national perspective applicable at a local institution-specific level? Studying the local conditions of an institution could provide different results than national studies because of an institution's local context. Further, campuses are responsible for the cohort of students that borrowed to attend their institution and defaulted (Hillman, 2014). As such, individual campuses are interested in knowing which specific factors affect their students' ability to repay their loans. Since individual campuses are responsible for their default rates, an institution needs to know which factors are associated with their students' defaulting. A case study is necessary to develop such a campus-specific understanding. Therefore, my study is a quantitative institutional case study of student default at a commuter institution. It is justified to conduct a single institution study so that campus leaders can know which variables are associated with default for their students. To fill this gap, I studied loan default for a population of commuter students who attended a nonresidential university. I sought to discover the predictors of federal student loan default for this population and tried to understand if institutional data could improve estimates of default probabilities for these students. My research questions, therefore, are:

1) What are the characteristics of Commuter State students who do and do not default?

2) What predicts default among students at Commuter State?

3) Do institution-specific, student-level measures improve estimates of default for students at Commuter State?

To offer context for my research questions, it is essential to understand research on: (a) student loans, (b) student loan default, and (c) commuter students. My study focuses on the

existing research of commuter students because my subject institution's student body is comprised only of commuter students. Connecting these three works of literature provide a unique context for my study within the existing research landscape. Following the discussion about student loan debt, default, and commuter students, I finish the chapter with a discussion of the purpose of my research and the significance of the study.

Student Loans

Student loans are necessary for many students to enroll and stay enrolled in higher education. As such, it is important to understand student loans as a mechanism to defer the cost of higher education enrollment, the levels of debt accrued by borrowers, and the risks associated with student loans and defaulting.

A Mechanism to Defer the Cost of Higher Education Enrollment

The higher education system in the U.S. places a significant financial responsibility on students and their families to pay for tuition and fees (Heller & Callender, 2013). Average tuition and fees for public, 4-year institutions have increased by 37% compared to 2007, and 110% compared to 1997 (Baum et al., 2017). Students who enroll in higher education, on average, cannot afford the total cost of attendance at the time of enrollment (Scott-Clayton, 2018). As a result, students, and in some cases, their parents, seek financial aid to help subsidize and defer the cost of attendance. To close the gap between what students and families are expected to pay and the cost of attendance, students, in turn, rely on financial aid, including student loans, to help subsidize the cost. The two primary sources of financial aid funds are from the federal government and higher education institutions themselves (Baum et al., 2017). Typically the financial aid office of a student's college or university is the primary source of allocation for the aid.

There are two primary forms to award financial aid: aid that does need to be repaid and aid that does not. Scholarships and grants are common forms of student financial aid that do not need to be repaid. On the other hand, student loans are a form of assistance that does need to be repaid by the borrowing students. Scholarships and grants reduce the price of enrollment, while student loans are a mechanism to defer the cost of enrollment to a later date. The significance of student loans is highlighted by the prevalence of federal student loan borrowing.

The Levels of Debt Accrued by Borrowers

Nationally, student loan debt has reached record levels, reaching approximately $1.6 trillion in outstanding student debt (Federal Reserve Bank of New York, 2021). Student loans are a financing mechanism for the majority of students in the U.S. higher education system (Heller & Callender, 2013). In 2010, student loan debt surpassed credit card debt, making it the largest non-housing consumer debt in the U.S. (Mezza & Sommer, 2015). The percentage of undergraduate students taking federal subsidized and/or unsubsidized student loans increased from 29% in 2006–07 to 38% in 2011–12 and back down to 25% in 2020–21 (Ma & Pender, 2021). Although the proportion of students borrowing federal funds decreased since 2011, federal loans in 2016–17 remained the largest source of financial aid for undergraduate students, representing 32% of the total aid provided by all sources including federal, state, and institutional aid (Ma & Pender, 2021). Since 2016–17, federal loans are the second largest source of student aid for undergraduate students behind institutional grants (Ma & Pender, 2021). More than approximately $45 billion in federal loans were provided to undergraduate students in 2020–21 (Ma & Pender, 2021). The largest source of financial aid, institutional grant aid, was approximately $58 billion, representing 33% of the total aid dollars allocated to students (Ma & Pender, 2021).

In 1997, total undergraduate federal loan aid was roughly $30.6 billion and in 2007 loan aid was $47.1 billion (Baum et al., 2017). Loan aid reached its peak in 2010–11 at $77.5 billion (Baum et al., 2015) and has declined in the years following to its current level, in 2020–21, of approximately $45 billion (Ma & Pender, 2021). Researchers suggest that, the increase in total student loan debt was driven by the number of borrowers (Dynarski & Kreisman, 2013). Twenty-three million borrowers held student loan debt in 2004 compared to almost 43 million in 2021 (Federal Reserve Bank of New York, 2021).

Over time, the number of borrowers and the average amount borrowed per student has increased (Council of Economic Advisors, 2016; Dooney & Yannelis, 2015). Between 2000–01 and 2015–16, the percentage of 4-year nonprofit bachelor degree recipients who borrowed increased from 54% to 60% (Baum et al., 2017). From 2015–16 to 2019–20, the percentage decreased to 55% (Ma & Pender, 2021). In the same period, the average debt per borrower increased from $22,100 to $28,400 (when adjusted for inflation to 2020 dollars) (Ma & Pender, 2021). The number of borrowers and borrowing levels are connected; however, this does not address whether borrowers' characteristics are associated with levels of borrowing. Research suggests that borrowers' characteristics are associated with borrowing, and differences emerge when considering borrowers' socio-economic status (SES), age, dependent status, and race (Baum et al., 2017).

According to filing data from the Department of Education's Free Application for Student Aid (FAFSA)—the application students must complete for federal aid—the largest volume of borrowers and the most significant growth of borrowing occurred among the lowest income families, measured by family income from federal tax filings. Specifically, low-income families' borrowing spiked from 2009 to 2011 and, although it has declined since, in 2016 it

remained above borrowing of FAFSA income filers of an income of $30,000–$75,000 and

$75,000 and above (Council of Economic Advisors, 2016). This spike resulted in a higher

proportion of Pell-eligible students compared to previous years.[1] According to the Department of

Education, 32% of enrolled undergraduates received a Pell Grant in 2016–17, compared to only

24% in 2006–07(Baum et al., 2017). The value of the grant has significantly lagged behind the

rise in tuition and fees, thus creating a greater financial responsibility for low-income students.

The value of the Pell Grant, measured in 2017 constant dollars, has increased by only 44% and

17%, from 1997 and 2007, respectively, compared to 110% and 37% increases in average tuition

and fees for public, 4-year institutions in the same time periods (Baum et al., 2017). In

summation, the rise in average tuition, the increase in low-income student enrollment, and the

lagging value of the Pell Grant have resulted in higher costs passed along to students and their

families.

Beyond measures of students' SES, debt levels vary with age. In 2014, for graduates 23

or younger, 34% accumulated no debt, and 11% accumulated $40,000 or more in debt.[2] For both

categories of graduates 24 to 29 and 30 to 39, 21% acquired no debt, and 25% and 33%,

respectively, accumulated $40,000 or more in debt (Baum et al., 2015). The remaining borrowers

were within the $1 to $39,999 range of debt. The data suggest the older the student, the higher

the likelihood they will use student loans to finance some (or all) of their educational expenses.

Further, the older the student, the higher the probability they will accumulate more than $40,000

in debt.

---

[1] Federal Pell Grants are "awarded to undergraduate students who have exceptional financial need" (Federal Student Aid, 2018a, p. 1).

[2] The analysis presented in Baum et al. *Trends in Student Aid 2015* segments debt into six categories including: (1) no debt, (2) less than $10,000, (3) $10,000 to $19,999, (4) $20,000 to $29,999, (5) $30,000 to $39,999, and (6) $40,000 or more.

Debt also varies based on dependent/independent student status.[3] A higher proportion of dependent students acquired no debt (34%) compared to their independent peers without dependents (25%) and with dependents (23%) (Baum et al., 2015). Not only were independent students more likely to accumulate student debt but a greater proportion of independent students' accumulated debt in the highest debt category ($40,000 or more) (Baum et al., 2015). Eleven percent of dependent students assumed $40,000 or more in debt compared to 25% and 29% for independent graduates without dependents and independent graduates with dependents, respectively (Baum et al., 2015).

Similar to a students' age and dependent/independent status, differences exist in debt levels when considering race. For 2011–12 bachelor's degree recipients, Asians were the least likely to acquire student loan debt (Baum et al., 2015). Forty-three percent of Asians accrued no debt, compared to 32% of White degree recipients, 27% of Hispanics, and 14% of Blacks (Baum et al., 2015). The amount of debt accrued varied by race as well. In 2011–12, 32% of Black bachelor's degree recipients accrued $40,000 or more in student debt, compared to 7% of Asian graduates, 16% of Whites, and 17% of Hispanics (Baum et al., 2015). As the data suggest, borrowing has reached record levels, students with a wide range of varying characteristics acquire student loan debt, and differences emerge in the levels of borrowing based on various student characteristics. In the next section, I discuss student loan default and begin to make the connections between borrowers' characteristics and defaulting.

---

[3] According to the Office of Federal Student Aid website, "An independent student is one of the following: at least 24 years old, married, a graduate or professional student, a veteran, a member of the armed forces, an orphan, a ward of court, or someone with legal dependents other than a spouse, an emancipated minor or someone who is homeless or at risk of becoming homeless" (Federal Student Aid, 2018b). All other students are considered dependent (Federal Student Aid, 2018b).

The Risks Associated with Student Loans and Defaulting

Student loans and the likelihood of default pose risks to individual borrowers and society. The prevalence of borrowing and the average amounts borrowed has garnered national attention. Hillman (2015) points out that media outlets routinely refer to the "student debt crisis" (p. 36) or the "student loan bubble" (p. 36). Akers and Chingos (2016), however, argue there is no evidence of a "wide-spread, systemic student loan crisis" (p. 4), but, rather, a narrative that largely focuses on anecdotes and "inappropriate framing of the issue" (p. 4). Existing research has not substantiated the "bubble" and "crisis" labels; however, scholars have identified problem areas regarding student borrowing that need further investigation (Akers & Chingos, 2016; Hillman, 2015). The following discussion examines these bifurcated risks.

Student loans pose risks to individual borrowers. Loan amounts could affect an individual's ability to access additional credit for purchases such as housing and cars (Hillman, 2014). Outstanding balances jeopardize an individual's credit to debt ratios, adversely affecting credit scores, which can impact the individual's ability to borrow money in the future. Federal loan debt is not eligible for bankruptcy and, therefore, a borrower's debt will follow them regardless of their financial situation (Akers & Chingos, 2016; Hillman, 2014).

Beyond the implications of student loans associated with individuals, there are associated risks to society as well. The U.S. government has allocated billions of dollars to the student loan program (Baum et al., 2017) which makes student loan repayment not an isolated responsibility for individual borrowers, but also a subject with considerable implications nationally. The Federal Direct Loan program for undergraduate students, which is administered and financially backed by the U.S. government, operates at a cost to taxpayers (Delisle, 2016). National headlines have painted a different picture. The typical national headline suggests that the

9

government earns revenue from the student loan program (Delisle, 2016). Overall, the Federal Direct Loan program does make money; however, when segmenting the loan program into two categories: (a) undergraduate loans and (b) graduate student and parent loans, the data reveal the undergraduate loan program costs the government billions of dollars while the graduate and parent loan program earns billions of dollars (Delisle, 2016). The Congressional Budget Office (CBO) estimated the cohorts of undergraduate loans from 2016 to 2026 will cost the government, and ultimately taxpayers, $19.6 billion (Delisle, 2016). There are two particular reasons that undergraduate loans are a cost to the government: (a) a subset of undergraduate loans are interest-free while students are enrolled, so the government is paying the interest accruing on the loan instead of the student; and (b) undergraduates are about three times more likely to default compared to other student loan borrowers such as graduate students and/or parents of students (Delisle, 2016).

Student Loan Default

The current undergraduate student loan situation does pose a substantial financial risk to both individual borrowers and the U.S. as a whole (Delisle, 2016). As such, it is vital to ensure that borrowers can repay their loan obligations. To do this, it is essential to understand the predictors of student default to minimize the risk of borrowers' defaulting on their loans. Once we establish a better understanding of defaulting, policies can be developed to help ensure student borrowers can repay their loans. The following discussion describes what student loan default is, how it is measured, and who is most likely to default.

What is Student Loan Default

Borrowers of student loans enter into repayment upon exiting higher education, with or without an earned degree. To be categorized in good standing, borrowers must remain current on

their repayment (Department of Education, 2018a). Conversely, borrowers who do not consistently repay their loans become classified as delinquent. If delinquency lasts for 270 days or more, borrowers enter into a status of default (Department of Education, 2018a). The two main ways to get out of default, beyond repaying the defaulted loan in full, are loan rehabilitation or loan consolidation (Department of Education, 2019a). While this default description refers to the individual, it is essential to also have an understanding of the aggregated picture of borrower default.

How Loan Default is Measured

To understand the levels of student loan default, the Department of Education aggregates defaulted borrowers based on the institution they attended while borrowing the loans and the year they entered repayment. The aggregation provides a default cohort and an institution-level measurement of the percentage of borrowers defaulting. The Department of Education captures default in a 3-year window once repayment begins (Department of Education, 2019a). Default rates are reported via 3-year cohort default rates (CDR), also known as a 3-year CDR (Department of Education, 2019c). Previously, a 2-year CDR was calculated; however, starting with the 2009 cohort, the Department of Education changed the measurement to a 3-year rate to create a more comprehensive snapshot of defaulting. These calculations are used to understand the magnitude of defaulting based on populations of students for individual institutions and higher education sectors.

Who is Most Likely to Default

Understanding the triggers of student loan default is an important topic (Gross et al., 2009). The existing research on student loan debt suggests the strong predictive characteristics of defaulting are whether a student graduated and their background characteristics—those

characteristics a student brings with them before entering higher education (Gross et al., 2009). Significant differences in default emerge between graduates and non-graduates (Hillman, 2014; Looney & Yannelis, 2015). In 1999–2000, 5% of graduates and 11% of non-graduates defaulted on their student loans within two years of entering repayment (Baum et al., 2015). In 2005–06, 2-year default rates were 5% and 19% for graduates and non-graduates, respectively. For the 2011–12 cohort, 2-year default increased to 9% and 24%, respectively. When disaggregated by sector, 9% of all borrowers from public 4-year institutions defaulted within two years after repayment for the 2011–12 cohort; however, only 6% of graduates defaulted compared to 18% of non-graduates (Baum et al., 2015).

Beyond graduation, defaulting is associated with borrower background characteristics. Research suggests default rates depend more on student and institutional factors than the amount of debt a borrower incurred (Looney & Yannelis, 2015; Scott-Clayton, 2018). These student background characteristics include, but are not limited to, socio-economic status, race/ethnicity, parental education levels, and prior academic preparation (Blagg, 2018; Gross et al., 2009; Looney & Yannelis, 2015; Scott-Clayton, 2018). Student loan borrowers from low-income families are more likely to default than there more affluent peers (Herr & Burt, 2005; Steiner & Teszler, 2005). Borrowers of color, particularly Black students, have a higher likelihood of defaulting than their White peers (Blagg, 2018; Scott-Clayton, 2018). Beyond borrower background characteristics, the type of institution a borrower attends is also associated with defaulting. Borrowers attending for-profit schools are more likely to default compared to their public and private non-profit school peers (Hillman, 2014; Looney & Yannelis, 2015; Scott-Clayton, 2018). In the 2016 3-year cohort, 15.2% of for-profit borrowers defaulted compared to

9.6% and 6.6% of borrowers from public and private non-profit institutions, respectively (Department of Education, 2018b)[4].

While researchers of student default have studied a range of student groups, commuter students are one group of students not studied. Which is important to note because my study's population was exclusively commuter students. Before researching commuter student default patterns and characteristics, it is first important to understand the characteristics and experiences of this group of students. The next section discusses commuter students' unique features and experiences as part of the institutional context of my study.

Commuter Students

The U.S. higher education system serves a diverse student body. It is varied by race/ethnicity, age, student life situation, campus residential status, socioeconomic status, and parental educational attainment status, among many other factors. This diversity creates a plethora of ways to categorize students into different populations. Commuter students, for example, comprise a majority of undergraduate students, and yet there are many gaps in existing research about this population (Biddix, 2015; Jacoby & Garland, 2004; Melendez, 2019). Commuter students represent more than 85% of the student population in the United States across all institutional types (Gianoutsos, 2011), and enrollment trends suggest that the proportion of commuter students will continue to grow and become more diverse (Horn & Nevill, 2006; Jacoby & Garland, 2004). For this study, my research focused on this large yet under-researched group of students. Before trying to understand default and commuter students, which is the context at my subject institution, it is essential to understand the unique attributes of commuter students and how they are different from their residential peers.

---

[4] The 2016 3-year cohort was one of the last cohorts before the 2020 worldwide pandemic. For this reason, I do not discuss the subsequent 3-year cohorts as I do not know if the debt and default data were impacted by the pandemic.

Regardless of institutional type, commuter students represent a wide range of ages and racial backgrounds, as well as living arrangements (e.g., living with parents or by one's self) (Chickering, 1974; Jacoby, 2000; Keeling, 1999; Tinto, 1975, 1993; Wilmes & Quade, 1986). Further, commuter students face particular challenges as a result of their commuting status. For example, Biddix's (2015) review of research on commuter students from 2005 to 2015 found one of the core challenges for commuter students is establishing institutional identity and engaging in campus activities. These engagement and identity challenges arise from the competing non-academic life demands (such as work and family) that commuter students experience alongside their academic demands (Biddix, 2015). Many commuter students also find that they feel like strangers in the new world of college (Jacoby, 2015). These events lead to a unique set of experiences when compared to residential students.

Research suggests that commuter students are fundamentally different from residential students (Chickering, 1974; Jacoby, 2000; Keeling, 1999; Tinto, 1975, 1993; Wilmes & Quade, 1986). For example, commuter students are more likely to have multiple life roles (e.g., parenting, full-time employment, community roles) compared to their residential student peers (Chickering, 1974; Jacoby, 2000; Keeling, 1999; Tinto, 1975, 1993; Wilmes & Quade, 1986). Commuter students not only have more heterogeneous background characteristics from their residential peers, but they also face different levels of campus engagement and experiences compared to their residential peers (Biddix, 2015; Mayhew et al., 2016). Contrast this with Mayhew et al.'s (2016) review of the current literature for students living on campus that found residential students experienced greater social and academic integration compared to their non-residential peers. Also, in an analysis of data from the National Survey of Student Engagement (NSSE), Kuh, Gonyea, and Palmer (2001) found that both first-year and senior students who

lived on campus reported more interactions with faculty members and higher levels of enriching educational experiences compared to their commuting peers. Ultimately, while researchers' have worked hard to understand the commuter population, the unique needs of this group have been neither adequately understood nor appropriately incorporated into policies, programs, and practices (Biddix, 2015; Jacoby & Garland, 2004). In order to ensure success for these students, more research is needed.

Purpose and Significance of the Study

Research suggests there are substantial economic benefits to earning a bachelor's degree (Council of Economic Advisors, 2016). In the first decades of the 2000s, more students are enrolling in higher education to capitalize on these benefits. The increased volume of enrolled students has resulted in an increasing number of students taking federal loans to defer the cost of attendance, and average levels of borrowing have steadily increased. In sum, more students are borrowing, and, on average, borrowing more. In the same timeframe, there have been increases in both the number of students entering repayment of their federal loans and the proportion of individuals defaulting on their repayments. As a result, researchers have studied student loan defaults to better understand this problem.

Existing default research has identified the important characteristics associated with default and this is important from a national context when trying to understand default. However, for institutions concerned with default for the students who borrowed to attend their campuses, more granular analysis is needed. The local context of the institution should be incorporated into understanding default for the institutions' borrowers. To establish a focused understanding of default at an institutional level, I studied students at a completely nonresidential institution, which I refer to with the pseudonym Commuter State. My approach allowed me to study the

association of the institution-specific characteristics of Commuter State to default. While the selection of my subject institution, Commuter State, addressed the issue of examining institution-specific characteristics, it introduced the challenge of sample bias. This sample bias means that in selecting this one specific institution, I gave up the ability to appropriately generalize across a large population of institutions. However, this selection did provide me with the opportunity to examine how institutional-level data can inform institutional decision-making about default. In addition, studying institutional-level data makes the information immediately useful for the subject institution to understand default characteristics. Given this context, in this quantitative study I asked three questions:

1) What are the characteristics of Commuter State students who do and do not default?

2) What predicts default among students at Commuter State?

3) Do institution-specific, student-level measures improve estimates of default for students at Commuter State?

By focusing on these questions I fill gaps in both the commuter student literature and the default literature. More specifically, the findings helped me examine specific default predictors for this group of students. This information could help inform the design of programs specific to the subject institution's default characteristics, maximizing the possibility of reducing default rates. In other words, by utilizing granular data that extends beyond administrative data, I hoped to gain a more specific understanding of the factors that predict default for an institutions' students and, thus, create targeted programs to address these shortcomings.

We know commuter students have a different college experience (Biddix, 2015; Mayhew et al., 2016), face different challenges, and are more heterogeneous based on many different characteristics when compared to residential students (Chickering, 1974; Jacoby, 2000; Keeling,

16

1999; Tinto, 1975, 1993; Wilmes & Quade, 1986). Existing research has not studied whether there is a link between loan default and the unique features and experiences of commuter students. My research examined the predictors of loan payment default that existing research has examined for other populations including variables such as student characteristics (e.g., race, age, gender, socio-economic status, first-generation status), graduation status, the field of study, and total federal loan amount. In doing so, the goal of my study was to expand existing understanding of loan default by incorporating institutional-level variables not traditionally captured in national datasets. By incorporating such variables, the model expands what is previously understood about default and provides findings that have practical implications. However, because my study examined commuter student from one institution, my findings are not generalizable across all commuter students at all institutions. The advantage of my study's institution specific approach is it allowed me to include institution-level data into the models to help understand their contributions to defaulting. My approach allowed me to examine how institutional-level data can inform institutional decision making about default, and perhaps provide a process for other commuter-based institutions to examine the issue. Specifically, by incorporating these institutional-level variables in an organized way (i.e., by type of activity, or by year), my findings could inform how early in a student's career I could identify them as "at-risk" of default.

The study's significance is its unique contribution to the underdeveloped, yet vitally important, intersection of literature on commuter students and student loan default. The existing literature has identified the need to further understand commuter students (Baum, 2005; Biddix, 2015; Dugan et al., 2008; Jacoby & Garland, 2004; Kodama, 2002; Pascarella & Terenzini,

17

2005; Clark, 2006) and student loan default (Gross et al., 2009). By design, the results of my study provide practical changes non-residential colleges can take to mitigate student default.

Methodology Overview

To answer my research questions, I used both descriptive statistics and regression analyses. I utilized descriptive statistics to provide insights into whether there is anything substantively interesting about the population of defaulting students compared to the not defaulting students at the subject institution. The descriptive analysis results helped answer research question one.

Next, I utilized logistic regression to identify predictors of default for commuter students at a nonresidential master's comprehensive. The findings of the initial regression analyses addressed research question two and helped me develop a model to predict default. Finally, my design utilized a block regression technique that *a priori* specified a sequence for adding sets of predictor variables to my model (Lomax & Hahs-Vaughn, 2012; Pallant, 2010). The order of the blocks into the model was important because I was trying to see if the predictive power of the models improved or leveled-off as I entered additional blocks. For example, my findings provide a model to know whether I could predict in their second year whether students would default as well as I could predict default in subsequent years, such as their third and fourth. The order of predictor variables' entry into the model is informed by previous research and is discussed in the conceptual framework section of chapter 3.

Summary

Research suggests there are substantial economic benefits to earning a bachelor's degree (Council of Economic Advisors, 2016). In the twenty-first century, more students are enrolling in higher education to capitalize on these economic benefits. The increased volume of enrolled

students has resulted in an increasing number of students taking federal loans to defer the cost of attendance, and average levels of borrowing have steadily increased. More students are enrolling in higher education, more students are borrowing, and borrowing levels are growing. In the same timeframe, there is an increase in the number of students entering repayment for their federal loans and an increase in the proportion of individuals defaulting on their repayments. Existing research seeks to understand the factors associated with defaulting; however, the current research also cites more work is necessary for this field of study. A way to further investigate this field of study is to research default at a local, institution-specific level. Therefore, I sought to understand if institutional data can improve estimates of default probabilities for this population at a nonresidential campus. In the following chapter, I discuss the existing literature on my topic, providing an understanding of the current gaps in this field of knowledge and, thus, situating my study within the existing research. In chapter 3, I outline the methodological approach I utilized for my quantitative case study of student loan default at a commuter institution. I present my research results in chapter 4. I conclude my study with chapter 5 which provides a summary of my findings as well as a discussion of the applicability of my research findings to the subject institution and beyond.

CHAPTER 2: REVIEW OF LITERATURE

Student loans are a common means of funding education for many students attending college in the United States and, consequently, repaying college debt has become the new norm of early to middle adult life borrowers (Williams, 2013). Although there is a great deal of literature on student loans, debt, and default, less is known about the intersection of these topics for specific populations, including commuter students and how local institution-specific measures can contribute to our understanding of default. With that in mind, in this study I asked:

1) What are the characteristics of Commuter State students who do and do not default?

2) What predicts default among students at Commuter State?

3) Do institution-specific, student-level measures improve estimates of default for students at Commuter State?

In chapter one, I provided context regarding student loans, student loan default, and commuter students. I discussed the connectedness of these three fields of knowledge and shared the importance of studying this intersection further. To situate the contribution my research can make to existing knowledge, in my literature review, I further discuss research related to my study (Creswell, 2009). In the review, I relate my study to the larger literatures on student loans, student loan default, and commuter students, ultimately demonstrating how my study fills gaps and extends prior studies (Cooper, 1984; Marshall & Rossman, 2006). In this chapter, I, then, provides "a framework for establishing the importance of the study as well as a benchmark for comparing the results with other findings" (Creswell, 2006, p. 25). To provide the context needed for my study, in this chapter, I first provide a brief synthesis of the financing of higher education in the U.S., discussing how the U.S. higher education system evolved into a loan-driven structure. Next, I examine student loan debt reviewing what the literature indicates about

the association between loans and factors such as students' likelihood to enroll, likelihood to persist, and likelihood to successfully transition to life beyond higher education. After reviewing the loan research, I discuss the findings of loan default, including what factors the research suggests associates with the likelihood to default. Finally, since my study is a case study of a nonresidential campus, it is important to understand what research says about commuter students. Therefore, I conclude my reviewed literature with a discussion of this specific population. Following the reviewed literature, I discuss my approach to conceptualizing default. My study conceptualizes default by linking the concepts from the previous literature review sections together into a comprehensive framework that informs my study. The chapter ends with a summary of how the existing findings of previous research inform the development and direction of my study. I discuss how my research project is unique and contributes to the existing knowledge about default and, more specifically, if institutional data could improve estimates of default probabilities for a nonresidential commuter campus.

A Brief Synthesis of the Financing of Higher Education in the United States

It is important to understand the financial structure of higher education in the U.S., and how the structure came to be in order to understand the current student loan debt and default situation in the U.S. The historical and current context of the funding of higher education helps explain how the financing of higher education in the U.S. resulted in the increased reliance on students and their families' to pay for their education. In the following discussion, I explain the sources of financing for higher education, including a brief synthesis of the historical policies and perceptions that contributed to the current structure.

The U.S. higher education system has six sources of financing that collectively pay the cost of student enrollment (Kane, 1999). These financial sources include federal grant and loan

programs, state and local government subsidies, institutional aid, university revenue generated from endowment income, tax relief programs, and, finally, students and families themselves (Kane, 1999). Although various sources exist, I primarily focus on the federal financing of higher education because of the significant size of federal aid, the vast growth in this aid, and because my research focuses on federal student loan default. Federal and state funding are not mutually exclusive of one another (Scott-Clayton, 2017); however, a shift in one source can affect the other. As such, to adequately understand the current status of federal loan debt and default, I included some state policy shifts within my discussion.

The federal government has always been the largest provider of direct aid to students (Scott-Clayton, 2017). State and local governments provide the next largest source of total support; however, much of this aid is via institutional appropriations ($73.5 billion in 2013–14 compared to $10 billion in direct grants to students) (Scott-Clayton, 2017). Not only does the federal government have the largest role in financial aid, but this role has also grown substantially over time. Since the late 1990s and early 2000s, the federal government has nearly quintupled its investment in financial aid, while state and local appropriations have increased by a mere 6% in the same period (Scott-Clayton, 2017).

The federal government's role in student financial aid was established in 1965 when President Johnson signed the Higher Education Act (HEA) into law (Scott-Clayton, 2017). The act instituted provisions for federal grants, loans, and work-study assistance (Scott-Clayton, 2017). The HEA of 1965 established the first government-backed loans known as the Guaranteed Student Loan program (Heller, 2008). These elements remain the foundation of undergraduate aid for college students (Scott-Clayton, 2017); however, changes in federal policy have caused significant shifts in the sources of aid available to undergraduates (Heller, 2013; Pascarella &

Terenzini, 2005; Scott-Clayton, 2017; Williams, 2013). In Heller's introduction to the book *Student Financing of Higher Education* (2013), he attributes four factors to driving the federal financial aid policy shifts: (a) a push for the massification of higher education with the intent to grow participation rates, (b) macroeconomic factors that lead to constraints on overall government revenues, (c) political factors which create competing demands for funding other services instead of higher education, and (d) a belief that the value-added from higher education is an individual gain and, thus, students should bear more of the burden to pay for it. As a result of these factors, greater emphasis has been placed on federal student loans, moving away from grants (Heller, 2013; Hillman, 2015; Pascarella & Terenzini, 2005; Scott-Clayton, 2017; Williams, 2013).

Student loans have long been prominent features of financial aid packages (Pascarella & Terenzini, 2005); however, factors such as the beliefs mentioned above about higher education have contributed to significant shifts in federal financial aid policies from grants to loans. The federal 1992 Reauthorization of the Higher Education Act outlined new loan program rules (Pascarella & Terenzini, 2005). The reauthorization increased the limit of money students could borrow annually and cumulatively in the federal loan programs (Heller, 2008). The shift from grants to loans passed the responsibility of financing higher education from taxpayers to students and their families (Heller, 2013; Scott-Clayton, 2017). The reauthorization also "liberalized the needs-testing that students underwent to qualify for the loans" (Heller, 2008, p. 40). This change, and the introduction of unsubsidized loans, led to an increase in the number of borrowers (Heller, 2008), paving the way to the current national debt situation.

Beyond federal measures, state policy shifts have affected student loan debt. Hillman, for example, (2015) indicates states' divestment in public higher education support is a contributing

factor to increasing debt levels. The U.S. Government Accountability Office (GAO) found, "persistent state budget constraints have limited funding for public colleges" (GAO, 2014, p. GAO highlights). From fiscal years 2003 through 2012, state funding for all public colleges decreased, while tuition rose (GAO, 2014). Such trends, in turn, increase the policy pressure for expanding federal financial aid (Scott-Clayton, 2017).

Compared to the late 1990s, more students are receiving more aid and more types of aid (Scott-Clayton, 2017). In 2013–14, full-time undergraduates received 50% more total aid compared to students in 2003–04 (Scott-Clayton, 2017). This increase in aid represents a mix of grants, federal loans, other assistance, and tax credits (Baum et al., 2014). Scott-Clayton (2017) argues, "the stakes have never been higher to ensure the effectiveness of financial aid—not just for the sake of the stakeholders who provide it but for the sake of students themselves, who make the biggest investments of all" (p. 2). Building on this understanding of the current funding of higher education students in the U.S., in the next section, I discuss student loan debt in the U.S.

Student Loan Debt

As the data and research suggest, student loan borrowing has grown significantly since the late 1990s and early 2000s. The previous section discussed how perceptions and polices contributed to this growth; however, it is also important to address how students are affected by these changes. In this section, I review the research that addresses this issue. Authors have examined student loans in many ways, including how they relate to enrollment, persistence, and postsecondary outcomes. As my work focuses on federal loan debt, I begin this section with an outline of the federal student loan program and provide supporting data to contextualize the significance of the program. Following the discussion of the federal loan program, my review of the loan research tracks students' chronological enrollment pattern. The review begins with

research looking at loans and students' decision to enroll. The next section reviews the role of loans and student persistence. The review concludes with the research on outcomes, particularly the association of loans and major decisions and loans and future earnings. Finally, I discuss the studies that look at other borrowing factors, particularly factors that do not align as neatly into one of the enrollment pattern categories mentioned above.

Overall, the debt research reviewed for this section is informed by Cho et al.'s (2015) review of student loan literature as well as other instrumental research. Cho et al. (2015) synthesized existing student loan literature and concluded that the use of student loans affects many individuals and households in the United States. The following discussion provides context for the federal student loan program and supports our understanding of the existing loan research.

The Federal Student Loan Program

The primary federal student loan program, and the group of loans I focus on in this study, is the Stafford Loan program. The program provides two types of loans for undergraduates: subsidized and unsubsidized loans (Scott-Clayton, 2017). Subsidized loans do not accrue interest while students are enrolled in higher education and are available only to students with financial need (Department of Education, 2019b). Unsubsidized loans accrue interest while students are enrolled and are available regardless of financial need (Department of Education, 2019b). In 2020–21, federal student loans accounted for 26% of all undergraduate aid distributed to students (Ma & Pender, 2021). During that same year, the largest source of aid provided to undergraduates were institutional grants, representing 33% of the total aid offered (Ma & Pender, 2021). Further, total federal loans increased by 134% in the ten years between 2000–01 and 2010–11, and increased 62% in the twenty years between 2000–01 and 2020–21, , but declined

by 40% between 2011–12 and 2020–21 (Ma & Pender, 2021). Federal subsidized and unsubsidized loans peaked in 2010–11 at over $104 billion in inflation-adjusted 2020 dollars. The figure was over $62 billion for 2020–21 but roughly only $44 billion two decades ago in 2000–01 (Ma & Pender, 2021).

Many students utilize the Stafford Loan program; it has grown since the 1990s and has a sizeable amount of federal dollars tied to the program. Given the scope and significance of the program, it is important to understand the interconnectedness between loans and student borrowers. In the following sections, I discuss the research on student loans and how the findings inform my study.

Student Loans and Enrollment

Scholars have studied the role student loans play in students' decisions to enroll in higher education. A major problem in understanding this relationship between loans and enrollment is the fact that loans have become an important essential part of financing college for many students as a result of grant aid no longer being sufficient to offset total out-of-pocket costs (Heller, 2008). There was a time when many students were able to finance their postsecondary education themselves (Heller, 2008). In this prior era, loans were commonly a vehicle to help finance a more expensive private education (Heller, 2008). In 2021, many students rely on loans to pay for college, regardless of the cost of the institution (Ma & Pender, 2021). For instance, it is even common for students to take out loans to attend relatively low-priced community colleges (Heller, 2008; Ma & Pender, 2021).

Utilizing a randomized experiment technique, Booij, Leuven, and Oosterbeek (2012) investigated the role that college students' knowledge and information played in their decision to enroll in college. The study found the information presented to students about student loans,

including details such as loan conditions, interest rates, and repayment periods, did not significantly influence students' enrollment decisions (Booij et al., 2012). This could be partially explained by the findings of Chudry et al. (2011) in that when considering undergraduate students' borrowing attitudes, students considered education loans as a way to enhance their future, rather than a form of debt. The study found that parents contribute to helping shape students' attitudes towards debt (Chudry et al., 2011). However, when accounting for students' racial and ethnic backgrounds, researchers found differences in the relationship between student loans as they relate to enrollment. Perna (2000) looked at differences in financial aid sensitivity among students from different racial groups. She found that the use of student loans to defer the cost of attendance reduced the probability of African American students enrolling (Perna, 2000). Perna (2000) concluded this might be due to an aversion of borrowing or an expectation that future earnings will be insufficient to repay the loans. Burdman's (2005) research expanded Perna's findings to low-income and minority families. Utilizing interviews conducted with students, counselors, and financial aid directors, Burdman (2005) found the need to borrow money for college impedes some students, particularly those from low-income and minority families, citing it is a barrier for items such as a lack of loan literacy, loan aversion, and lack of confidence in their ability to repay their debt obligation (Burdman, 2005). Research considering students and families from various racial and ethnic backgrounds found debt aversion was a result of inadequate knowledge about financial aid and an expectation that future earnings would be insufficient to repay the loans (Burdman, 2005; Heller, 2008; Perna, 2000). Even if the characteristics above are not barriers to loans, cultural differences across racial groups can affect individuals' willingness to incur debt (Burdman, 2005; Heller, 2008; Perna, 2000)

Scholars have also looked at how the financial aid process for student loans can impact enrollment. Johnson's (2012) qualitative study of students and parents engaging in the FAFSA (Free Application for Federal Student Aid) process found that students were less engaged in the details of the borrowing compared to their parents. Johnson's research suggested student borrowers did not read correspondence and forms as much as parents; they did not know the terms of their loans, and some did not even know the amount they received (2012). Monks (2012) found that the types of financial aid policies at institutions can be important to who applies for admission, enrolls, and the level of debt accrued. For example, Monks (2012) found that need-blind admission policies increased the probability of college enrollment of low-income students, and also increased the average level of student debt overall.

In this section, I highlighted studies linking student loans and enrollment. This research is essential for my study because it helps provide context for the enrolled students at the subject institution, including background context about why specific populations borrowed and how much they borrowed. In the following discussion, I transition to the effects student loans have on students' college persistence.

Student Loans and Persistence

Scholars have studied the impact of loans on college persistence. Hu and St. John (2001), for example, examined what types of student financial aid packages were effective in promoting persistence. They researched the mix of the sources of financial aid in aid packages but did not look at the specific amounts of each form of aid. They found that student loans, when packaged with grants, had a positive impact on persistence (Hu & St. John, 2001). Their research found financial aid packages with a mix of loans and grants had a particularly positive impact for

students from different racial groups (Hu & St. John, 2001), with the largest impact on Hispanic and African American students, as compared to White students (Hu & St. John, 2001).

Beyond the mix of different forms of student aid, scholars have found the amount of aid matters. Paulsen and St. John (2002) found borrowing was negatively related to student persistence, particularly for low-income students, concluding that the problem was the total aid allocated was inadequate to help the students meet their college expenses. Johnson's (2012) research examined specific levels of borrowing. The author concluded allowing students to borrow up to the full cost of attendance each year raised college completion by only 2.4%, compared to 5.3% if tuition subsidies (such as grants) were provided. These findings suggest the mix, type, and amount of financial aid matters for completion. The impact of loans to completion is connected to other forms of aid (such as grants and scholarships) received by a student. This is important for my study because students in my population received a similar mix of grants, scholarships, and loans. In order to understand the federal student loan default, it is important to account for the findings mentioned above. My study accounts for the previously mentioned findings by including the different types and amounts of aid (such as university grants and scholarships) allocated to students at my subject institution.

In addition to loans, grants, and scholarships, parental resources are other common means to fund an undergraduate education. Like loans, parental resources are a means to finance an education and can directly impact the amount of debt a student acquires. Keane and Wolpin (2001) examined the difference in college completion for high- and low-income families. Their research found that parental financial support for college explained differences in completion rates (Keane & Wolpin, 2001). The findings that parents' ability to pay is a greater contributor to a likelihood to graduate is an essential element of the interaction between the effects of wealth

and completion. The effect of parental resources is important to my study because it could relate to the amount of loans students acquired and their persistence. My study attempted to account for students' parental support by incorporating background economic measures into the model. This served as a proxy for parent's ability to pay; however, it is important to note just because parents' ability to pay is high doesn't mean their willingness to pay is the same.

Beyond the parental ability to pay, students' access to credit has been examined as a factor associated with graduating. The findings are mixed and suggest the type of credit access can have different associations with graduation. Keane and Wolpin (2001) found that credit availability had only a marginal effect on completion. Lovenheim's (2001) and Stinebrickner and Stinebrickner's (2008) research challenged these findings. Stinebrickner and Stinebrickner (2008) found that students with credit constraints were less likely to graduate. The students surveyed for this research revealed they would have borrowed more if credit was available (Stinebrickner & Stinebrickner, 2008). Lovenheim (2001) took the research beyond access to credit and looked at specific types of credit. The author found that college enrollment for low-income students increased if they had access to additional home equity credit (Lovenheim, 2001).

Cho et al. (2015) concluded from their literature review that borrowing constraints affect students' higher education decisions. However, the borrowing constraints are not strictly financial proxies. Other factors are contributing to the probability of college graduation and the students' willingness to borrow, particularly student background characteristics (Bound, Lovenheim, & Turner, 2010; Stinebrickner & Stinebrickner, 2008). Like the research I discussed in this section, my study design controlled for the effect of student characteristics.

Student Loans and Outcomes

While issues of persistence are often discussed in the context of student loans, another area of study focuses on student loans and other outcomes. For example, studies show that student loans affect outcomes such as choice of major and future earnings. Kuzma et al. (2010) showed that students' choice of major drove their confidence in debt management for undergraduate business students at a public 4-year institution. The study concluded that junior- and senior-level business students' confidence in their ability to secure employment and manage debt was significantly related to their debt levels. Another study examined the outcomes of a university that replaced the loan component of financial aid awards with grants (Rothstein & Rouse, 2011). The authors found that debt caused more graduates to choose higher-salary jobs rather than lower-paying ones (Rothstein & Rouse, 2011). After the university's shift to no debt, more students graduated from majors leading to lower salary careers (Rothstein & Rouse, 2011). These findings suggest that debt is associated with students' decisions about major. These studies help inform my study design by providing evidence about why I should include major of study as a variable within my model.

Beyond the choice of major, outcomes research examined the relationship between student loans and wealth accumulation. The findings suggest wealth can be adversely affected by debt burdens. Using a panel of national data, Elliot and Nam (2013) found that living in a household with student debt in 2009 was associated with having $40,000 less in assets compared to living in a household with no student loan debt. Further, household net worth was inversely related to outstanding student loan debt; however, households with a 4-year graduate, regardless of debt level, had higher net worth than households without a 4-year graduate (Elliot & Nam, 2013). In another study of the impact of student loan debt on lifetime wealth, Hiltonsmith (2013)

calculated that dual-headed households with bachelor's degrees from 4-year universities with an average debt level for the study's time period led to a lifetime wealth loss of over $200,000. Households with greater than average student loan debt levels, all else equal, were projected to have even larger wealth loss over a lifetime (Hiltonsmith, 2013). These studies help inform the potentially unmeasured factors in my model's design. Knowing the association of these unmeasured factors and debt helps with understanding the findings of my model analysis.

Student loan debt can affect health and transition to adulthood. A 2004 study examined the impact of debt and the effect of attitudinal measures on student mental health of students in their final year of study reflecting on their university experience (Cooke, Barkham, Audin, Bradley, & Davy, 2004). The researchers found that students' perceptions of their finances and debt were associated with their mental health scores (Cooke et al., 2004). Students with serious financial concerns reported feeling more tense, anxious, or nervous, among other factors, compared to their peers with low financial concerns (Cooke et al., 2004). A similarly focused study in 2013 researched students' financial anxiety and debt, which included their student loans (Archuleta, Dale, & Spann, 2013). Factors reported to be connected with financial anxiety included financial satisfaction, student loans, and gender (Archuleta et al., 2013). Several researchers have utilized national datasets to analyze the relationship between debt and transitions to adulthood (Cho et al., 2015). Key findings across the studies suggest that student loans lower the likelihood of both marriage and becoming a parent (Cho et al., 2015). These findings inform the multitude of factors beyond academics that students are thinking about while enrolled in college and beyond. Knowing these factors exist, and their connectedness to debt, helps with the interpretation of my study's findings, providing additional context that is not

directly in the research design. These unmeasured factors could be useful when analyzing the results of my model.

Student debt affects millions of Americans. In fact, according to the most recent data from the Federal Reserve, it affects 44 million people with an outstanding balance of approximately $1.6 trillion (Federal Reserve Bank of New York, 2021). Existing research suggests student debt can affect factors including students' decision to enroll in higher education, persistence to graduation, health, the transition to adulthood, and accumulation of wealth. Some of these factors adversely affect borrowers, and some positively affect borrowers. Borrowing to defer college costs to a later date helps provide students the opportunity to engage in higher education; however, upon exiting higher education, with a degree or not, repaying the debt is the second phase. Not all students repay their loans on-time or at all. The following section delves into the existing research about student loan default.

Student Loan Default

The existing literature on student loan default considers three sets of factors/variables that impact student loan default: (1) pre-college, (2) in college, and (3) post-college. As such, this section of the literature review discusses the findings within these categories of factors/variables. I begin with a discussion of *pre-college* characteristics as they relate to defaulting. Pre-college variables include race/ethnicity, age, gender, family structure, parental education, income, and academic preparation. The next section discusses the *in college* variables that relate to students' college experiences, including factors such as enrollment patterns and program of study. I conclude the student loan default section with a review of the findings of *post-college* characteristics as they relate to defaulting. This set of variables includes debt burden, educational attainment, and characteristics of the economy such as unemployment rates.

Gross et al.'s (2009) thorough review of student loan default research largely influences my review of the literature for this section. The authors' study reviewed only the strongest methodologically sound studies, is well organized by thematic topics, provided clear and concise generalizations of research findings up to 2009, and many of the student loan default studies published after 2009 referenced Gross et al.'s (2009) study. Overall, student loan default research strongly suggests student characteristics and background are significant predictors of loan default (Dynarski, 1994; Flint, 1997; Hillman, 2014; Knapp & Seaks, 1992; Looney & Yannelis, 2015; Monteverde, 2000; Podgursky et al., 2002; Scott-Clayton, 2018; Steiner & Teszler, 2003, 2005; Volkwein & Szelest, 1995; Volkwein & Cabrera, 1998; Wilms et al., 1987; Woo, 2002). The following discussion provides more detail about the association between the factors mentioned above and defaulting.

Pre-College Characteristics

Ethnicity is perhaps the most studied characteristic in the loan default literature (Gross et al., 2009). Research findings have been consistent in that students of color are more likely to default than their White peers (Volkwein & Szelest, 1995; Woo, 2002). Specifically, African Americans were found to be at greatest risk of defaulting (Knapp & Seaks, 1992; Podgursky et al., 2002; Steiner & Teszler, 2003; Wilms et al., 1987) and were less likely to resume repayment after defaulting compared to their White and Asian American peers (Volkwein et al., 1998). Dynarski (1994) found the relationship of ethnicity and defaulting to be statistically significant regardless of institutional type. The understanding of why students of color are more likely to default compared to their peers is not well understood. Authors have suggested the effect is due to emergent differences in borrowing levels, family finances, levels of satisfaction after graduation, and employment trends (Wilms et al., 1987; Volkwein et al., 1998). An explanation

for African American default may lie in their employment and earnings after college. Authors document African Americans experience economic marginality by factors such as earning less and having higher rates of unemployment compared to their White peers (Wilms et al., 1987). This marginality can lead to lower lifetime accumulated wealth and higher levels of dissatisfaction with their educational experience, which can adversely affect their ability and willingness to repay their loans (Wilms et al., 1987). Wilms et al. (1987) also suggest there are additional and unmeasured economic variables that are associated with default. Volkwien et al.'s (1998) nationally representative sample of student loan borrowers discusses a set of characteristics that may contribute to the unmeasured economic variables mentioned above. The authors found that African Americans and Hispanics in their sample had almost twice the number of dependent children and almost twice the rate of separation and divorce (Volkwein et al., 1998).

Studies focusing on age nearly all found that as age increases, so does the likelihood of defaulting (Podgursky et al., 2002; Looney & Yannelis, 2015; Steiner & Teszler, 2005; Woo, 2002). As age increases, so can greater potential family responsibilities, additional accrued levels of debt (beyond student loans), and additional job responsibilities, among many other possibilities. More recent studies challenged previous research concerning age and defaulting. Younger borrowers are at a far greater risk of defaulting and delinquency (Dynarski & Kreisman, 2013; Looney & Yannelis, 2015). In Cunningham and Kienzl's (2011) analysis of student default, 28% of students under 21 defaulted, compared to 18% of borrowers between the ages of 30 and 44, and 12% of those 44 and older.

The connection between gender and loan default, unlike age, is much less clear in the literature. Several studies found no significant difference in the likelihood of default between

men and women (Volkwein & Szelest, 1995; Wilms et al., 1987). Several other studies found that men were more likely to default than women (Flint, 1997; Podgursky et al., 2002; Woo, 2002). The differences in findings could be a result of the populations studied. The studies with no significant difference in the likelihood of default were nationally representative in contrast to Podgursky et al.'s (2002) and Woo's (2002) samples that were state-level studies in Missouri and California, respectively.

Examining other characteristics, scholars have found academic preparation—defined as high school rank, high school GPA, and standardized test scores—contributes to loan default significantly. As high school rank, GPA and standardized test scores increase, the likelihood of defaulting decreases (Podgursky et al., 2002; Steiner & Teszler, 2003; Woo, 2002). Woo (2002) found that one standard deviation increase in cumulative high school grade point average (half a grade in this case) reduced the borrower's chances of defaulting by nearly 14%. Interestingly, Podgursky et al. (2002) showed that increases in the ACT composite score reduced the likelihood to default. However, as students progressed to their degree (measured by continuous semesters of enrollment), the association between ACT composite score and default became insignificant.

Family structure can mitigate or contribute to defaulting, depending on the dynamics of the family structure. For example, the greater the number of dependents claimed by a student, the greater the likelihood of loan default (Dynarski, 1994; Volkwein & Szelest, 1995; Woo, 2002). Further, Volkwein et al. (1998) found that being a single parent was a significant contributor to the likelihood of default. Conversely, students who had a family safety net, such as parental support, were less likely to default than their peers who had no family support (Looney &

Yannelis, 2015; Volkwein et al., 1998; Woo, 2002). Looney and Yannelis (2015) found that dependent students were less likely to default compared to their independent classmates.

Parental education is another significant indicator of student default (Volkwein et al., 1998; Volkwein & Szelest, 1995). Students with parents with higher levels of education were less likely to default compared to their first-generation college peers. Steiner and Teszler (2003, 2005) found the parental interaction associates with both mother's and father's education levels. The association with default holds whether one considers the mother's or the father's level of education, and so it is not about which parent has the highest level of education, but rather the parents' overall highest level of education.

The literature regarding family income and student default suggests, pre-college, the higher the family income, the greater the socioeconomic status of the students' family, the less likely it is a student will default (Hylands, 2014; Knapp & Seaks, 1992; Looney & Yannelis, 2015; Mezza & Sommer, 2015; Wilms et al., 1987; Woo, 2002). Post-graduation—different from departing college pre-degree—the greater the earned income of the student the less likely the chances of defaulting (Dynarski, 1994; Looney & Yannelis, 2015; Volkwein et al., 1998; Woo, 2002). Conversely, unemployment is strongly associated with defaulting (Dynarski, 1994; Monteverde, 2000).

The student debt burden has an inverse relationship to the likelihood of defaulting (Dynarski & Kreisman, 2013; Hyland, 2014; Mezza & Sommer, 2015). The lower the debt amount, the greater the likelihood to default. Receiving a degree helps explain the inverse relationship between borrowing levels and default. Students who graduated and borrowed were more likely to have higher levels of loan debt than their borrowing colleagues who dropped out (Dynarski & Kreisman, 2013; Hyland, 2014; Mezza & Sommer, 2015). These authors suggest

debt levels are tied to time in college, such that a borrower who graduates is in college longer and is likely to take on more debt that a student who drops out. Moreover, recent studies concluded that student loan balances are generally not a significant predictor of student loan delinquency (Dynarski & Kreisman, 2013; Hyland, 2014; Mezza & Sommer, 2015). Both Dynarski and Kreisman (2013) and Hyland (2014) found that borrowers with lower levels of debt defaulted at the highest rates. The average loan amount in default is $14,000 compared to the average loan amount in good repayment status, $22,000 (Dynarski & Kreisman, 2013). Taken from a different perspective, 16% of borrowers defaulted, while only 11% of total loan dollars are in default (Dynarski & Kreisman, 2013). Using current data, in March 2021, 17% of borrowers were in default which represented 11% of outstanding federal student loan dollars (Ma & Pender, 2021). Defaulters have lower average debt levels than borrowers who successfully repay their loans (Ma & Pender, 2021). In the second quarter of 2021, defaulters owed an average of $21,700, compared to $35,400 for those in repayment (Ma & Pender, 2021).

In College

Scholars have also focused on the association of students' college experiences and loan defaults. The variables discussed include enrollment patterns, the program of study, and persistence. Progress toward a degree is a significant predictor of not defaulting (Podgursky et al., 2002). The greater the number of consecutive semesters a student is enrolled, the less likely they are to default regardless of graduation status (2002).

Another factor, the students' program of study, appears to affect the likelihood of defaulting in two ways: (a) amount of debt incurred, and (b) post-graduation earnings. Harrast's (2004) research, which reviewed one institution's students' debt as it related to their major, found a list of majors that were likely to contribute to higher levels of debt compared to the rest

of the institution's programs; however, the author did not know why some majors resulted in greater debt burdens. Other scholars have found that post-graduation earnings related to the field of study affect personal income and, therefore, one's ability to repay (Flint, 1997; Herr & Burt, 2005; Steiner & Teszler, 2005; Volkwein et al., 1998; Volkwein & Szelest, 1995). For example, Volkwein et al. (1995) found that science or technology majors were incrementally less likely to default compared to their other peers from other majors. Herr and Burt's (2005) regression model found the school of the students' degree to be a predictor of default; however, individual majors were not. These findings are somewhat mixed but do indicate that major and college of study are important to include in my default model.

The single strongest predictor of not defaulting, regardless of institution type, is postsecondary degree completion (Dynarski, 1994; Knapp & Seaks, 1992; Looney & Yannelis, 2015; Mezza & Sommer, 2015; Volkwein et al., 1998; Woo, 2002). Researchers' findings spanning three decades, from the 1990s to the second decade of the 2000s, consistently conclude that graduating is strongly linked to not defaulting. Dynarski (1994) and Volkwein et al. (1998) both analyzed the National Postsecondary Student Aid Study (NPSAS-87) data, a nationally representative sample of borrowers who left postsecondary school from 1976–85. The authors of both studies found individuals were more likely to default if they exited without a degree than their graduated colleagues (Dynarski, 1994; Volkwein et al., 1998). Moving into the early 2000s, Steiner and Teszler's (2005) research, which studied Texas postsecondary students who entered repayment during federal fiscal years 1997, 1998, and 1999, estimated students who graduated had a 2% chance of defaulting compared to 14% for those who did not graduate. Woo (2002) studied Californian borrowers who took out student loans in the federal fiscal year 1995. The author connected borrower background information from the FAFSA with post-college

employment data from the state of California (Woo, 2002). Woo (2002) found that leaving college without a degree was a significant determinant of defaulting. Mezza and Sommer (2015) analyzed a nationally representative dataset of individuals spanning from 1997 through 2010 that included credit bureau records, FAFSA background data, loan information, college enrollment and completion records, and school characteristics. Their research found, like their colleagues before them, degree attainment is a significant predictor of not defaulting (Mezza & Sommer, 2015). In the discussion of their findings, Mezza and Sommer (2015) discuss default is not driven by large levels of debt, but rather, by factors correlated to the ability to repay it. As such, students who graduate have great earnings potentials (Looney & Yannelis, 2015) and have higher associated credit scores (Mezza & Sommer, 2015), which are associated factors with the ability to repay student loan obligations. In order to address my research questions, my study included some of the in-college variables discussed. Specifically, graduation is in my model due to its important association with defaulting.

Post-College

Thus far, I have presented the literature regarding the associations between pre-college student characteristics and defaulting and in-college experiences and defaulting. The third type of student borrowers' experiences that can affect their ability to repay their loans occur post-college. For example, scholars have studied the health of the economy at the point of students' labor force entry. Looney and Yannelis (2015) argue that the Great Recession had a substantial effect on educational enrollment and borrowing. A poor economy (a) decreases the opportunity cost of college enrollment, and (b) puts downward pressure on financial metrics such as earnings and asset valuation (Bound, Lovenheim & Turner, 2010), increasing higher education enrollment across all sectors (Looney & Yannelis, 2015). Increased pressure on borrowing coincided with

40

increasing enrollments. During the Great Recession, state budgets were cut for many public institutions, and access to alternative credit markets was restricted, affecting both students' ability to borrow as well as their parents' ability to borrow to finance their dependent's education (Looney & Yannelis, 2015). The higher education sector realized an increase of more than 2 million additional borrowers per year from 2009 to 2011 compared to 2003 to 2007 (Looney & Yannelis, 2015). The deflated Great Recession economy realized suppressed employment opportunities as a result of company hiring freezes, potentially adversely affecting the labor market for student borrower graduates and dropouts of the higher education system. Akers and Chingos (2016) showed an increase in default rates during and after the Great Recession, which is consistent with the notion that default increases during economic recessions and decreases during periods of growth. Looney and Yannelis (2015) found that the poor post-college economy for students who exited higher education during the Great Recession was associated with defaulting for some populations of students but not all students. One such sub-group of students are students of color who are more likely to be unemployed compared to their White peers and, thus, could affect their ability to repay their loans (Volkwein et al., 1998). Post-college factors are essential to understanding default. However, access to this information, such as tax records that provide employment status and income, are challenging to attain, and linking the health of the economy to the granular impact to individuals is outside the resources I have available for my study. As such, my study omitted the post-college factors discussed.

In this section, I identified relevant student loan default literature and discussed its' findings. Overall, the default literature has identified the pre-college, in college, and post-college student characteristics, experiences, and economic conditions associated with defaulting. Now

that I have provided foundational information about the variables associated with default, I

discuss in the following section the research on commuter students.

Commuter Students

Commuter State, the institution for my case study, is exclusively a commuter campus. In

order to understand default for Commuter State students, it is important to understand the context

of the students attending the institution: commuter students. In this section, I discuss the existing

research findings of this group of students. It is important to note, because practitioners and

scholars' overall understanding of commuter students is lacking, there is a need for further

research about this population (Biddix, 2015; Melendez, 2019), a comprehensive review of the

existing literature of commuter students conducted in 2015 by Biddix, coupled with my review

of the literature, guides the following discussion about commuter students. It includes a

discussion of how commuter students are defined and the number of commuter students enrolled

in higher education, the characteristics of commuter students, their experiences while enrolled in

higher education, and their success rates measured by retention and graduation.

Commuter Students Defined

In this section, I begin with a discussion of the different ways commuter students are

defined, discuss how my study defined commuter students, and conclude with a discussion of the

size of the commuter population. The largest group of students in the U.S. higher education

system are commuter students (Biddix, 2015; Melendez, 2019); however, the classification of

these students occurs in a multitude of ways. I share two particular ways the students can be

classified, which aligns with the National Survey for Student Engagement (NSSE) (NSSE,

2017). The first definition includes students who have lived on campus their freshman year and

the following years live off-campus in nonresidential housing within close proximity to campus

with fellow university students. This represents a traditional student's university experience at a predominantly residential campus. Contrast the traditional experience to the commuter experience whereby the student commutes to campus throughout their entire collegiate career, living with their parents, family members, or commuting from their residence as an independent student. My research adopts the latter classification of commuter students which aligns with Newbold et al.'s (2006) definition of a commuter student, stating "a commuter student is defined as one who does not live on campus but attends the university from local and surrounding areas" (p. 142). I extend the definition further, however, to those students who never lived on-campus. I accomplish this by utilizing a completely nonresidential university as the site for the study. In the following section, I discuss this population as it relates to enrollment in U.S. higher education.

To understand a national perspective of the proportion of students who commute compared to those who live on-campus, I obtained data from NSSE, which provides a comprehensive assessment of first-year and senior students' residential situations. NSSE includes a question regarding students' housing options: "Which of the following best describes where you are living now while attending college?" (NSSE, 2017). The survey response options include: (a) dormitory or other campus housing, (b) residence, walking distance, (c) residence, driving distance, (d) fraternity or sorority house, or (e) none of the above (NSSE, 2017). The following is the aggregate response of enrolled first-year students' residential status at the subject institution: 8.2% dormitory or other campus housing,[5] 13.4% walking distance, 70.1% driving distance, 0% fraternity or sorority, and 8.2% none of the above (NSSE, 2017). Comparing this to all NSSE participants in 2017, a greater portion of students nationally are

_____

[5] The subject university does not have "dormitory or other campus housing." I hypothesize these respondents were students who lived in an apartment building across from campus that is privately owned. If this is the case these responses would increase the "walking distance" category.

living in campus housing as first-year students than are commuting. For the 2017 NSSE survey administration, 62% of respondents reported living in dormitories or other campus housing, 7% in walking distance, 25% in driving distance, 1% in a fraternity or sorority house, and 5% indicated none of the above (NSSE, 2017). For the 2017 survey administration, the results for senior students represent the expected migration of students out of campus housing and becoming "commuter students." For the subject institution, senior respondents reported 2% lived on campus, 6% in walking distance, 85% in driving distance, and 7% indicated none of the above. Reviewing national data, 13% of seniors reported living on campus, 23% in walking distance, 55% in driving distance, 1% in fraternities and sororities, and 8% indicated none of the above. As the data suggest, my subject institution includes only commuter students, those who were commuter students during their entire collegiate experience, a stark contrast to "commuter students" who lived in campus housing their first year and moved off campus into walkable and driving distance residence by their senior year. My subject university is not the only institution in the U.S. that serves students who are commuter students their entire collegiate career. The completely nonresidential setting of my subject university is an important and unique context that I had to consider to conduct my quantitative case study of defaulting at Commuter State.

As another way to understand the size of the U.S. higher education commuter student body and provide a context for the students of my case study, I review the prevalence of nonresidential universities in the U.S. The Carnegie Foundation developed a classification methodology to categorize higher education institutions. In 2005, Carnegie expanded its classification methodology which was based on three questions: "What is taught (Undergraduate and Graduate Instructional Program classifications)?, who are the students (Enrollment Profile and Undergraduate Profile)?, and what is the setting (Size & Setting classification)?" (Carnegie

Foundation Website, 2018). The Carnegie classification recognizes 2,576 4-year higher

education institutions (both for-profit and not-for-profit). Out of the 2,576 institutions, 1,151

(45%) institutions are 4-year primarily nonresidential (commuter) universities. Nonresidential

institutions have fewer than 25% of students living on campus or have more than 50% of their

students enrolled as a part-time status (Carnegie Foundation Website, 2018). These primarily

nonresidential institutions enrolled roughly 6.3 million or 46% of students in 2017. The data

mentioned above suggest just under half of all enrolled students at 4-year institutions are enrolled

at primarily nonresidential campuses, which suggests there are a sizeable number of institutions

enrolling a large number of students who could have similar commuter experiences as my study

population.

Commuter Student Characteristics

Thus far, I have discussed the unique context of the student body for my subject

institution. I outlined that Commuter State's students are exclusively nonresidential. In the

following section, I discuss the characteristics of commuter students. It is important to note, the

following discussion on commuter students refers to a broader definition of commuter students

than I have outlined in the previous section. This is the most relevant research relative to my

defined population and, therefore, is reviewed to provide an understanding of the population of

students.

Commuters are fundamentally different from residential students (Chickering, 1974;

Jacoby, 2000; Keeling, 1999; Tinto, 1975, 1993; Wilmes & Quade, 1986). They are more

heterogeneous compared to their residential peers, representing a wide range of ages, racial

backgrounds, socioeconomic groups, birth origin, and living arrangements (e.g., living with

parents or by one's self) (Chickering, 1974; Jacoby, 2000; Jacoby & Garland, 2004; Keeling,

1999; Melendez, 2016; Tinto, 1975, 1993; Torres, 2006; Wilmes & Quade, 1986). Commuter students are more likely to be older than 22, represent a higher proportion of underrepresented minority populations, be a non-U.S.-native, and have different living arrangements compared to their residential peers (Chickering, 1974; Jacoby, 2000; Jacoby & Garland, 2004; Keeling, 1999; Melendez, 2016; Tinto, 1975, 1993; Torres, 2006; Wilmes & Quade, 1986). Further, enrollment trends suggest that the proportion of commuter students will continue to grow and become more diverse (Horn & Nevill, 2006; Jacoby & Garland, 2004). Newbold et al. (2011) surveyed seniors at a mid-sized state university. Comparing the responses of commuter and non-commuter students, the authors found commuter students were more likely to (a) be non-traditional (measured by age), (b) be transfer students, (c) work more hours, (d) earn more income, (e) be less likely to be involved in school-sponsored activities, (f) be less likely to believe their university has a good reputation, and (g) be less likely to identify with the university (Newbold et al., 2011). Other authors have further identified unique characteristics of commuter students, finding these students are more likely to have multiple life roles (e.g., parenting, full-time employment, community roles) compared to their residential student peers (Chickering, 1974; Jacoby, 2000; Keeling, 1999; Tinto, 1975, 1993; Wilmes & Quade, 1986). In an analysis of the National Center for Education Statistics' profile of undergraduates in U.S. postsecondary institutions, 1999–2000, Jacoby and Garland (2004) found undergraduate students 24 years old or older were almost all commuter students. Commuter students are more likely to work off-campus, creating a "three-point commute between home, campus, and work" (Jacoby & Garland, 2004, p.71). Overlaying findings from the National Center for Education Statistics (2002) with previous findings, Jacoby and Garland (2004) found commuter students are more likely to work, to work more hours, and to work off-campus than residential students. As a result, commuter

46

students could be perceived as being less committed to their studies than residential students. However, Jacoby states, "the educational goals of commuter students are very similar to those of residential students" (2004, p.63), but "student" (Jacoby & Garland, 2004, p.63) may not be the primary identity for commuters. Multiple life roles are common for commuter students. Beyond their higher education and work responsibilities, commuters are more likely to manage households, including children, siblings, and relatives (Jacoby & Garland, 2004; Melendez 2019; Wilmes & Quade, 1986).

Commuter Student Experiences While Enrolled in Higher Education

Biddix's review of the last decade of research on commuter students found one of the core challenges for commuter students is establishing institutional identity and engaging in campus activities (2015). The author explains these challenges have "not shifted dramatically" (p. 1) in the last decade of research on commuter students (Biddix, 2015). Mayhew et al.'s (2016) review of the current literature for students living on campus found that residential students experienced greater social and academic integration compared to their nonresidential peers. Also, in an analysis of data from NSSE, Kuh, Gonyea, and Palmer (2001) found that both first-year and senior students who lived on campus reported more interactions with faculty members and higher levels of enriching educational experiences compared to their commuting peers. The authors further found that commuter students often lack the support of the campus environment—an established benchmark of effective educational practices of NSSE (Kuh et al., 2001).

The support networks for commuters predominantly exist off-campus, which includes support from individuals such as parents, siblings, partners, children, and coworkers (Jacoby, 2004). As a result, these students have to negotiate the responsibilities and time commitments

that come with higher education, but with individuals who may not have the same level of understanding of higher education as individuals on-campus (Jacoby, 2004). Jacoby (2015) utilized transition theory to help explain commuter students' experiences with integration. For example, commuter students may find navigating university processes and systems difficult, may lack an understanding of the opportunities available on campus, and may struggle to make college part of their already busy lives (Jacoby 2015). This can lead to these students feeling like strangers on campus. Students who feel marginal are also less likely to engage in college experiences that are associated with educational success, such as high-impact educational practices (Schlossberg et al., 1989) that increase fundamental engagement on campus (Kuh et al., 2007). Kuh et al. (2001) did find that although many commuter students' ability to engage was limited by work and family, they put forth as much effort as residential students in areas related to the classroom. The research suggests that academic integration is an important component for all students' connection to their success in college (Johnson, 1997; Tinto, 2006). Regression analysis results from Melendez's (2019) research found that residential status was not a significant predictor of academic adjustment after accounting for gender and race or ethnicity. However, Melendez's (2019) study did indicate that college adjustment affects nonresidential and residential students differently in the way in which they integrate themselves into the social structures of the university, in particular with regards to social factors such as taking part in campus activities and meeting new people (Crede & Niehorster, 2012).

Commuter Student Success Rates

Researchers have studied the success rates of commuter students measured by both retention and graduation. Ishitani and Reid (2015), analyzing a nationally representative dataset, found there were no significant differences in first-year dropout behaviors between on-campus

and off-campus students. The authors' study segmented students' off-campus living into two categories, living with and without their parents. However, differences emerged when looking only at the segment of students who lived with their parents, as this off-campus population was roughly 23% more likely than on-campus students to drop out (Ishitani & Reid, 2015). Ishitani and Reid (2015) suggest that students living with their parents may spend less time engaging in academic and social activities on campus compared to their residential peers. Interestingly, when incorporating levels of academic and social integration into understanding retention for commuter students living with their parents, the findings shifted. As social and academic integration increased, the risk of departure due to living with parents diminished (Ishitani & Reid, 2015).

Beyond life circumstances and the difference in student background characteristics, studies examine commuter student mental health. Research findings suggest that students' emotional health can contribute to their likelihood to persist. Astin (2001) found that commuting had negative effects on students' self-assessment of their emotional health, indicating commuting was associated with raised levels of stress.

While existing research seeks to understand the needs of commuter populations, such as understanding factors related to their mental health, the unique needs of this group are neither adequately understood nor appropriately incorporated into policies, programs, and practices (Biddix, 2015; Jacoby & Garland, 2004), and aggregate national data supports these authors' findings. For example, 4-year commuter colleges comprise nearly 32% of the total number of colleges in the United States and have typically reported higher baseline dropout rates for students than residential campuses one, two, and three years after students enter college (Weissberg & Owen, 2005). The causes of the increased dropout rates may link to the many

unique concerns associated with commuting beginning with transportation, weather, and driving costs as well as more common psychosocial concerns such as balancing work, school, and household obligations, and establishing support networks (Gefen, 2010; Jacoby & Garland, 2004). These trends in retention carryover to graduation statistics. In an analysis of enrolled undergraduates, Astin (2001) found that commuting was negatively related to bachelor degree attainment and graduate school enrollment.

The chapter, thus far, outlined a brief history of the financing of higher education, followed by a discussion of what is known about student loan debt and student loan default. The literature review concluded with a discussion of what is collectively understood about commuter students within higher education. In the final section of the chapter, I outline the conceptual framework of the study.

Conceptualizing Default

According to Miles and Huberman (1994), the conceptual framework of a study "lays out the key factors, constructs, or variables, and presumes relationships among them" (p. 440). It is "a network of interlinked concepts that together provide a comprehensive understanding of a phenomenon or phenomena" (Jabareen, 2009, p. 51). Some studies link to an existing conceptual framework, and some studies are conceptualized outside of a pre-existing framework. My study is the latter, in that I am drawing from the thinking of numerous existing frameworks to inform my work. As such, I assemble the assumptions, models, and findings of existing research discussed in this chapter, including scholarship on student loan debt, loan default, and commuter students, combined with the following discussion of a student persistence theory, to guide my conceptualization of default. Through the synthesis of the existing literature, I devised a tentative conceptual framework for understanding default. In particular, the reviewed literature informs

my study's design, including the variables I used in my models and analyses. These details are

discussed further in Chapter 3. The following discussion explains how student persistence

contributed to my conceptualization of default. I conclude this section with an explanation of my

devised framework.

The research on student persistence informed my work. Persistence relates to default

because students who enroll but do not earn a degree are more likely to default than their peers

with a degree (Gross et al., 2009; Looney & Yannelis, 2015; Scott-Clayton, 2018). My study,

therefore, extends beyond the existing default research by digging deeper into what happens

during students' enrollment that may predict loan repayment, independent of persistence, and

may be more within the sphere of the university to affect. To expand the current understanding

of default, then, my empirical analysis is guided by Braxton et al.'s (2004) theory of student

departure in commuter colleges and universities. The authors developed the theory to address the

distinct differences between residential and commuter institutions (Braxton et al., 2004). For

example, commuter colleges and universities lack well-defined and structured social

communities compared to residential institutions (Braxton et al., 2004). Further, commuter

students face conflicting obligations to work, college, and family (Tinto, 1993). Given these

differences, the theory provides a framework for understanding student departure from commuter

institutions (Braxton et al., 2004).

Braxton et al.'s (2004) departure theory has four basic elements including student entry

characteristics, the external environment, the campus environment, and the academic

communities of the institution (Braxton et al., 2004). These elements directly influence students'

commitment to an institution or departure decisions (Braxton et al., 2004). Specifically, the four

elements are connected to students' initial institutional commitment, their subsequent

institutional commitment, and, ultimately, their persistence (Braxton et al., 2004).

The first element of the theory, student entry characteristics, are the characteristics with

which students enter into college. The theory's entry characteristics are the same as the pre-

college characteristics discussed earlier in this literature chapter. At commuter institutions, these

characteristics "play a significant role in the student departure process" (Braxton et al., 2004, p.

43). The second element of the theory is the external environment. It captures students' ability to

adjust to both the external environment and the environment of the institution (Braxton et al.,

2004). Commuter students frequently have obligations distinct from attending college (Tinto,

1993). These obligations are connected to a students' ability to adjust to their college experience,

which is connected to their institutional commitment (Braxton et al., 2004). The campus

environment, the third element of the theory, is connected to departure. The commuter campus

environment can be chaotic with minimal structured social communities. Commuter students,

many of whom have external priorities, typically spend minimal time on campus beyond their

coursework (Braxton et al., 2004). The transient, come and go, nature of the commuter campus is

connected to students' decisions of whether to depart or persist. The final element of the

departure theory is the academic communities of the institution. With the absence of a well-

defined social structure at commuter universities, the students' academic experience plays a

meaningful role in the student departure process (Braxton et al., 2004). The academic domain is

characterized by structured co-curricular activities or academic courses taught by faculty (Gross

et al., 2015). Mainly the classroom community is an influential component of student departure

decisions (Braxton et al., 2004). Participating in a learning community offered by faculty who

engage students in active learning experiences increases academic integration, can result in

greater academic outcomes, and decrease the likelihood of departure (Braxton et al., 2004). To understand if defaulting can be better predicted and predicted earlier, the academic domains, and thus the students' college experience, were modeled in my study. I, therefore, included factors that capture the academic domains to unpack defaulting further. The idea was to see if local data can improve current default models and, ultimately, help institutions improve student default rates. Tying together Braxton et al.'s theory of student departure in commuter colleges and universities with the existing default research frameworks helps to extend the existing body of literature.

My conceptualization of default is constructed into three blocks of variables, including pre-college student characteristics, college experiences, and post-attendance. My purpose of creating the blocks of variables, conceptually, was to understand the relationship of each block of variables to defaulting. Therefore, my conceptualization is the relationship of pre-college student characteristics to defaulting, the relationship of college experience to defaulting, and the relationship of post-attendance variables to defaulting. Previous default research findings primarily inform the pre-college student characteristic and post-attendance blocks of the model. These two blocks include the variables that the previous default research finds to be associated with default. Braxton et al.'s theory of student departure primarily informs the framework's college experience block. Each block includes different themes of variables as they relate to existing research findings and conceptual frameworks. The pre-college student characteristics block contains variables related to students' demographics (i.e., sex, race/ethnicity, age, and first-generation status), socio-economic status, and incoming academic characteristics (including incoming GPA and standardized test scores). The college experience block contains the academic domain variables identified by Braxton et al.'s departure theory (i.e., term GPA, term

credits attempted, and term credits earned). The third block, post-attendance, includes variables to capture student graduation and total loan amount. The connection of each of these three blocks of variables to default creates my conceptualization of default. It is important to note, relationships exist between the blocks (i.e., the relationship between pre-college student characteristics and college experiences, college experiences and post-attendance, pre-college characteristics and post-attendance variables) and these inter-relationships likely affected the estimates of the variables in my current analyses; however, these inter-relationships can be observed by looking at the stability of the coefficients in the model as the new blocks were added. Where the point estimates and statistical significance are stable across models, there is likely to be very little inter-correlation. Where those estimates and statistical significance change across models, I am likely to have inter-correlations that effect my interpretations of these estimates, so any inferences from these changing coefficients must be made cautiously or not at all. Chapter 3 provides further details on the specific variables included in each block and explains the order in which each category was loaded into the empirical model.

Summary

In 2009, a meta-analysis of the student loan default literature conducted by Gross et al. (2009) argued that the chief limitation of the existing research literature was its lack of current studies. Since Gross et al.'s (2009) literature review, researchers have published more statistically robust student loan default literature. However, a common theme for the recent articles is their call for further work in this area. Current student loan default research has mostly supported the findings of previous studies and has continued to call for more research in this area of study (Blagg, 2018; Hillman, 2014; Looney & Yannelis, 2015; Scott-Clayton, 2018). My

research further contributes to our understanding of student loan default by investigating how

local institution-specific data could improve estimates of default probabilities.

CHAPTER 3: METHODOLOGY

The purpose of this chapter is to detail the methodological approach I utilized for my study on commuter student loan default (Belcher, 2009; Creswell, 2009). Specifically, I discuss how I addressed my three research questions:

1) What are the characteristics of Commuter State students who do and do not default?

2) What predicts default among students at Commuter State?

3) Do institution-specific, student-level measures improve estimates of default for students at Commuter State?

I focused on these questions to extend the default research. Further, I structured my research questions to capitalize on findings that can inform institutional decision making and policies with a goal of identifying an approach to data analysis that leadership at individual institutions can replicate on their campuses.

In this chapter, I begin with a discussion of the study's sample/population: commuter students at a nonresidential campus. Then I discuss my data sources and procedures, including the process of accessing the data from multiple sources and merging it into one data set. After I establish the data creation process, I discuss the variables in the study, identifying the dependent and independent variables. The analytical method section follows the variables section to provide a roadmap for the analytic process, connecting the research questions specific to the quantitative methods utilized in the study. I conclude the chapter with an explanation of the limitations of the research.

Population and Sampling

As discussed in previous chapters, my quantitative case study's population was commuter students. Specifically, I studied commuter students who commuted to campus for their entire

collegiate experience. My purposeful selection of Commuter State as the case study institution

introduced the challenge of sample bias. This sample bias means that in selecting this one

specific institution, I gave up the ability to appropriately generalize across a large population of

commuter students. However, this selection did provided me with the opportunity to examine

how institutional-level data can inform institutional decision-making about default, and perhaps

provided a process for other commuter-based institutions to examine the issue. In addition,

studying institutional-level data makes the information immediately useful for Commuter State

to understand default characteristics and engage in change that can directly impact the likelihood

of defaulting for its students.

Focused on a completely nonresidential university, my study's population consists of

students from a medium-sized nonresidential regional master's comprehensive university located

in the Midwest—Commuter State. The subject institution is different from other institutions.

When comparing Commuter State to its 139 peer institutions within the Basic Carnegie

classification of Master's Colleges and Universities with Larger Programs and primarily

nonresidential campuses, differences emerge.[6] The student population at Commuter State is

different from its peers with regard to age and the diversity of enrolled undergraduates.

Commuter State's undergraduate population had a greater proportion of its student aged 18 to 24,

at 77%, compared to 53% for the entire set of primarily nonresidential campuses (Author's

calculation from NCES data). Although Commuter State has a younger enrolled student

population compared to the average of its peer cohort, the proportion of its undergraduate

---

[6] The Carnegie classification of Master's College and Universities with Larger Programs was based on the number of master's degrees awarded in 2016–17. Institutions that awarded at least 200 degrees were included among the larger programs category (Carnegie Foundation, 2018).

population that identify as White was greater than the average for the set of primarily

nonresidential campuses at 68% compared to 42%.

Other differences exist between Commuter State and its Basic Carnegie classified peers

for graduation outcomes. For the most recently reported cohort, 53% of Commuter State's Pell

recipients graduated within 6-years compared to a 36% average for the primarily nonresidential

campuses (Author's calculation from NCES data). Interestingly, Commuter State, on average,

has a similar proportion of Pell-Eligible students (44%) compared to their primarily

nonresidential peer institutions; however, Commuter State is able to graduate their Pell-eligible

students at a higher rate. When moving from just Pell recipients to the entire undergraduate

population, the success rates of Commuter State compared to its peers fluctuates. Twenty-two

percent of Commuter State's students graduated in 4-years, which is similar to the 21% for the

primarily nonresidential institutions (Author's calculation from NCES data). Commuter State's

5-year graduation rate of 48% is greater than the average for the primarily nonresidential

institutions at 36% (Author's calculation from NCES data). The success of Commuter State's

students continues to increase for the 6-year graduation rate. Commuter State's 6-year graduation

rate of 56% is 15% higher than the average for the primarily nonresidential campuses (Author's

calculation from NCES data). Because of the differences of Commuter State's population when

compared to peer institutions, my findings are not generalizable to the other primarily

nonresidential institutions. However, by studying one institution, and thus, incorporating local

data into my analyses, I have better data, meaning it is more specific, compared to national

studies. Therefore, my findings are directly applicable to Commuter State's understanding of

their defaulted students. Further, my findings for Commuter State may provide a model that

other nonresidential institutions can replicate on their campuses to understand loan default of their commuter students.

It is important to note that the population of Commuter State students is also the sample for the study. I utilized total population sampling, which is a type of purposive sampling technique, for my study because I am examining the entire population of students at Commuter State. Total population sampling is defined by choosing to examine an entire population that have a particular set of characteristics (Lund Research Ltd, 2012). In this case, the population examined were students from a nonresidential students, thus, Commuter State students. The particular set of characteristics were the students were strictly nonresidential for their entire collegiate experience while enrolled at Commuter State. The advantage of studying the entire population is that, whatever the results, no inferences are necessary, these are the results. If I want to interpret statistical significance it can be taken to a meta-population such as studying all student who attended nonresidential institutions. The key idea with population samples is that the differences between groups don't need to be statistically significant to be true differences for the population. Based on available data for students within the federal government's 3-year cohort default rate (CDR) metric, I focus on eight cohorts of students from 2009–16. When pooled, this totals 14,260 borrowers, 828 of whom defaulted within their respective 3-year timeframe for an aggregate default rate of 5.8% (see Table 1). The 2011 cohort has the highest 3-year CDR at 8.1%, and three different cohorts have the lowest at 4.6%. The 2016 cohort, the most recent, has a default rate of 6.0%.

Table 1. *3-Year Cohort Default Rate Repayment and Default Data*

Borrower Cohort	Number of Borrowers in Repayment	Number of Borrowers in Default	Three-Year Cohort Default Rate
2009	1,339	61	4.6%
2010	1,422	110	7.7%
2011	1,681	137	8.1%
2012	1,999	115	5.8%
2013	2,026	94	4.6%
2014	2,048	115	5.6%
2015	1,933	88	4.6%
2016	1,812	108	6.0%
Total	14,260	828	5.8%

The national default rate for students at all types of institutions in 2016 was 10.1%, representing 458,687 defaulted students, while the default rate is 6.8% (119,117 students) for students who attended 4-year public institutions (Department of Education, 2019d). These approximately 120,000 defaulted students attended many different types of public 4-year institutions ranging from flagship research institutions to small regional universities (Department of Education, 2019d). They enrolled at institutions ranging in degree offerings, enrollment sizes, and student SES. The institutions these students attended also varied from being primarily residential campuses to exclusively nonresidential. Due to the range of types of institutions, residential experiences, and types of students served in the public 4-year sector, the default rates are variable and do not provide much value for comparison purposes. To address the wide variation of 4-year institutional types and to connect the above data to my study, I calculated the average cohort default rates for institutions by campus residency. It is important to note, the previously discussed data presented default rates by individuals across institutional sectors. Due to limitations of data availability, Figure 1 represents the average CDR's for three different cohort years across all institutional types. Figure 1 shows that the average default rate for primarily nonresidential institutions is higher than both primarily residential and highly residential institutional type averages. Interestingly, the default rate for Commuter State is closer to the average for highly residential institutions than either the primarily nonresidential average or the primarily residential average. This suggests that, in some ways, Commuter State is different from the average commuter institution emphasizing the importance of local data for local decisions, and further supporting my purposeful selection of Commuter State for my study. Local data may help inform Commuter States' financial aid programs, where affordability is one of the university's top priorities for its students. Local findings may also provide additional

information to the annual review of the effectiveness of the university's financial aid budget model. It is not the intent for the Commuter State findings to be generalizable across the nation. Rather, the findings can help inform Commuter State leadership and possibly provide an approach for other institutions that seek to understand the triggers to default for their student bodies.

Figure 1. *Average 3-Year CDR by Institution Residency*
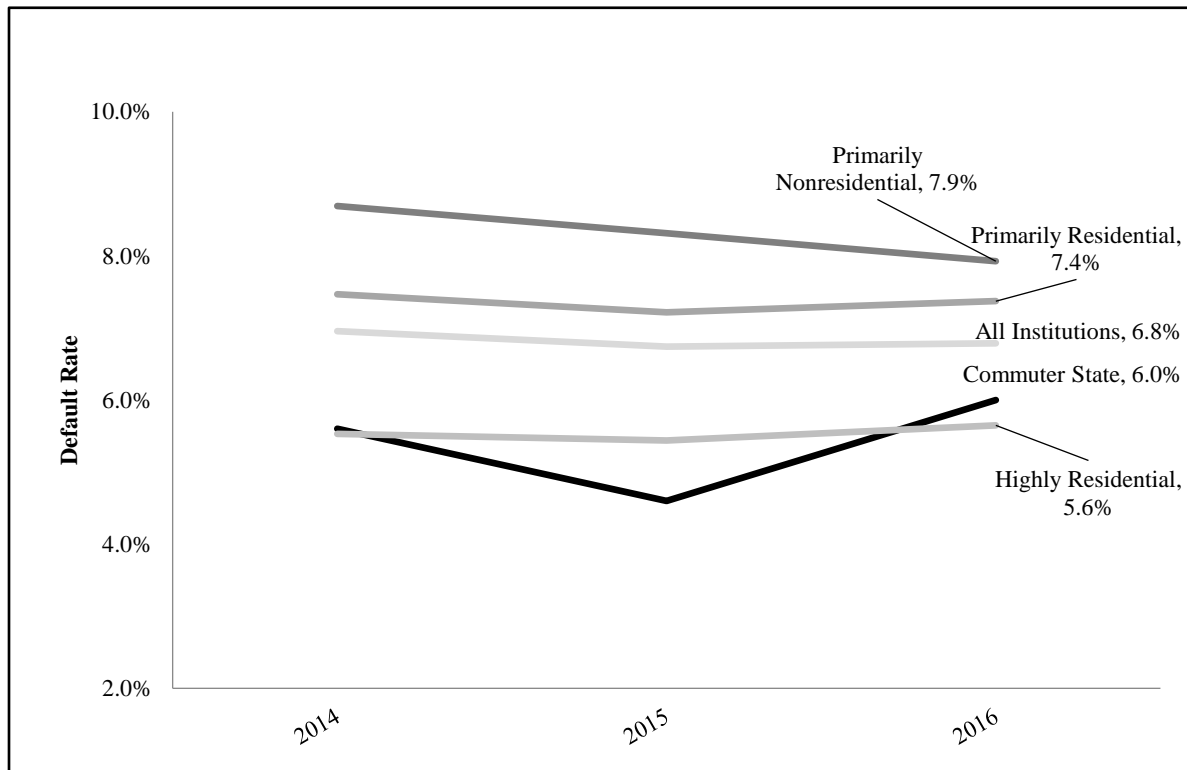
Overall, at Commuter State (and nationally), the vast majority of students are repaying their loans. However, the number of borrowers and levels of outstanding debt, as cited in the first two chapters, suggests a need to examine default more closely to ensure institutional practice and policies are mitigating the likelihood of default for their students. Neither national administrative nor institutional data exist that examines the characteristics of Commuter State's exclusively nonresidential population. To date, Commuter State has not examined the characteristics of those who have defaulted. Examples of such characteristics include graduation status, students' level of need, major enrolled, and level of borrowing. This descriptive information, in addition to the modeling results from the study, advances the existing understanding of default at Commuter State and provides some guidance for confirming how national-level findings can be consistent (and in some cases inconsistent) with local institutional situations. Further, my approach to these data provide approaches for how other institutions may examine and use their local data for local decision making.

Data Sources and Procedures

In this section, I discuss my study's sources of data, the procedures for accessing the data, and the procedures for linking the multiple sources of data together. The data set for my research requires students' background characteristics, their year-over-year enrollment, financial aid data, and whether they graduated from the institution. This information is included to provide as complete a picture as possible of the student borrowers' at the subject institution. The data collection process is a multifaceted approach with data accessed from multiple sources.

Data Sources

Data for this study came from the National Student Loan Data System (NSLDS) and the subject institution's student information system (SIS). NSLDS is the U.S. Department of

Education's central database for student aid (NSLDS, 2020). It receives data from schools, servicing agencies, the Direct Loan program, and other Department of Education programs (NSLDS, 2020). The NSLDS database is the sole source of data for students' default status. It also includes the students' cohort, loan status (such as in default, not in default, deferred, in repayment), and federal student loan amounts for loans received from the subject institution. The data to access from NSLDS includes student unique identifiers, the students' cohort, as well as their repayment status, loan type, and loan amount. The unique student identifier, repayment status, and loan amount are key variables from this resource. The unique identifier is the variable to connect borrowing data to enrollment and background information. The repayment status indicates whether a student has defaulted, which is the dependent variable for the regression model. The total loan amount indicates how much money a student borrowed.

The other source of data for my study is from Commuter State's student information system (SIS). The SIS is the university's database for all student information, including admission data, financial aid data, enrollment data, and student characteristics related to student campus engagement. The SIS populates from multiple sources, including the Free Application for Federal Student Aid (FAFSA), a students' admissions application, and enrollment and academic data from each semester of enrollment. The FAFSA updates students' financial information each year, contingent on the student filing their information each year. The data from the FAFSA provide instrumental student background information, including students' SES, first-generation status, dependency, and whether the student has dependents. The SES variables are critical to understanding the students' ability to pay and provide a proxy for their Pell Grant status. The data from the students' FAFSA submission in their first term of enrollment are the data utilized in the model. Typically, the FAFSA variables mentioned above do not change that

much year-over-year. Therefore, to streamline the data processing, only the first FAFSA filing results were utilized in the study.

The admission application data is another source for student background characteristics. The application data provides the student's race/ethnicity, high school or previous institution transfer GPA at admission, and composite ACT score. The students' background remains static in the SIS and only changes if the student requests a change to their information. Enrollment information is updated every semester for every student in the database. The period for the data collection for each student in the study's data set is from the first term the student enrolled at the institution until the student exited. While I would like to include data reflecting the time after the student exited, it is not available. I discuss this further in the limitations section.

Procedures for Data Access and Data Set Creation

The subject university strictly enforces access rights to student-level data. To ensure data is accessed, stored, and utilized within the university's standards, an application for IRB approval was submitted to the IRB offices of both Commuter State and Michigan State University. Commuter State's financial aid office assisted in accessing the NSLDS data. The data obtained from NSLDS included student-level information. This data can only be extracted from the NSLDS data system by the appropriate personnel and approvals for student information at Commuter State. By utilizing NASFAA's decision tree, the financial aid office deemed the use of this FAFSA data adhered to the standards for data permission for the Department of Education and was, therefore, able to release the information to me.[7]

---

[7] To ensure financial aid data is used appropriately, The National Association of Student Financial Aid Administrators (NASFAA) provides a data-sharing decision tree that is reflective of the Department of Education's legislative language regarding FAFSA data sharing (NASFAA, 2019).

The remainder of the data comes from Commuter State's SIS. The financial aid data stored in the SIS was extracted and shared by a professional from the institution's financial aid office. A professional from Commuter State's institutional research office, again, only after proper permissions were granted, extracted the enrollment and admissions data.[8] Throughout the entire data collection process, the data was stored on a password-protected laptop that had a dual authentication login process. This level of data protection aligns with the same processes of Commuter State.

After the data extraction, the three separate data sets were merged into one to create the final data set for the study. The three data sets include the data extracted from the NSLDS database, the SIS data set with financial aid information, and the SIS data from the enrollment and admissions application information. The data sets were merged using the students' unique identifier that was included in each of the three data sets. After merging, I cleaned the data for the analysis. For a brief description of the data cleaning process, see Appendix A.

Variables

In this section, I discuss the details of the analytical method, including the data analysis and the modeling techniques. Before discussing the analysis process, it is important to discuss the variables utilized in the model, including discussions of both the dependent and independent variables. The dependent variable is defaulting on federal student loans. As noted in previous chapters, my study focused only on federal student loan repayment, excluding any other types of loans such as private or personal loans taken for additional financing of the borrowers' education and related expenses. After the discussion of the dependent variable, I introduce the independent

---

[8] The reason the data from the subject institution's SIS is extracted as two files is due to data access permissions at the subject institution. The access permissions are granted to the university's professional staff based on their roles. As such, the financial aid and enrollment and admissions application data are extracted by two different individuals with different data access permissions.

variables used to predict default. The discussion of the independent variables includes the rationale used to categorize the independent variables into "blocks" to address the key ideas introduced in research question 3.

Dependent Variable

The outcome variable for the model is defaulting on student loan repayments within three years of exiting the subject institution. The Department of Education calculates default and includes the information in the NSLDS. The default variable is a binary categorical variable, meaning it has one of two outcomes; in this case, the categories of defaulted or not defaulted. Subject to the limitation noted above (i.e., focusing exclusively on federal student loans), my research did not include defaulting that may occur with private or personal loans students incurred during enrollment or defaults that occurred after the 3-year period. The limitations section further explains these issues. However, this is the best data that currently exists on the phenomenon and is used by all scholars and agencies interested in tracking and understanding the issue.

Independent Variables

In this section, I first outline the rationale for the types of variables in the model organized by blocks, then discuss the individual variables within each block. I discuss the variables of particular interest to the study in greater detail as part of their respective block discussion. Table 2 provides a detailed list of the independent variables and how they were blocked into the model.

Table 2. *Regression Model Variables and their Associated Source, Stored Location, Type, and Value*

Block	Variable	Source	Stored In	Type	Values
Dependent Variable					
	Defaulted	NSLDS	NSLDS	Categorical	(Not Defaulted) Defaulted
Independent Variables					
Post-Attendance					
	Graduated	SIS	SIS	Categorical	(Graduated) No Degree from Subject Institution
	Total Loan Amount	SIS	SIS	Continuous	Scaled to $1000 Increments
Pre-College Student Characteristics					
	Sex	SIS	SIS	Categorical	(Female) Male- unknown not reported is omitted
	Race/Ethnicity	SIS	SIS	Categorical	(White), Hispanic, Black, Other URM
	Age (at initial enrollment)	SIS	SIS	Continuous	By Year
	Expected Family Contribution (EFC)	FAFSA	SIS	Continuous	Scaled to $1000 Increments
	Pell-Eligible	FAFSA	SIS	Categorical	(Non-Pell) Pell
	Commuter State Grant (Merit and Need-based)	SIS	SIS	Continuous	Scaled to $1000 Increments
	First-Generation Status	FAFSA	SIS	Categorical	(Not First-Generation) First-Generation
	Dependency Status	FAFSA	SIS	Categorical	(Dependent) Independent
	Student has Dependents	FAFSA	SIS	Categorical	(No Dependents) Has Dependents
	Incoming GPA (HS GPA or Transfer GPA)	SIS	SIS	Continuous	Rounded to .1 increments
	ACT Composite Test Score	SIS	SIS	Continuous	
College Experiences by Semester Enrolled (n+1)					
	Term GPA	SIS	SIS	Continuous	Rounded to .1 increments
	Credits Attempted by Term	SIS	SIS	Continuous	Scaled to integers
	Credits Completed by Term	SIS	SIS	Continuous	Scaled to integers
	Cumulative Credits Completed	SIS	SIS	Continuous	Scaled to integers
	Term Enrollment Indicator	SIS	SIS	Categorical	(Enrolled) Not Enrolled
	Major	SIS	SIS	Categorical	(Non-STEM) STEM

To align my study with previous research and the theory of student departure in commuter colleges and universities, the blocks were structured based on grouping variables in the following themes: pre-college, college experiences, and post-attendance. The pre-college characteristics block includes variables that are inherit to a student, in other words, the factors students come with to Commuter State. The college experiences block consists of in-college variables that capture the students' college experiences by semester enrolled, including term GPA and credits attempted and completed. The post-attendance block includes student characteristics after attendance, the most relevant of which is whether the student graduated from the institution. Overall the blocks were loaded into the models differently for different models for different reasons. For instance, in my first model, I loaded only two of the three blocks to address research question two. The blocks loaded were the pre-college characteristics and post-attendance blocks. I only loaded these two blocks for my first model because researchers have previously found these variables, pre-college characteristics and graduation, are the primary predictors of default (Gross et al., 2009; Hillman, 2014; Looney & Yannelis, 2015; Scott-Clayton, 2018) and by establishing them in the model first, I can identify the significant predictors of default among students at Commuter State—which explicitly addresses research question two. To address research question three, I loaded all three blocks of variables into the model and loaded them in sequential order of enrollment. Thus, I loaded the pre-college characteristics block, then college experience blocks, and finally post-attendance block. My purpose was to determine if any subsequent variables, in this case local data, substantively improved model fit *in the order that the institution would receive them*. For this model, I loaded the pre-college characteristic block first because it includes variables students come with to Commuter State. Second, I loaded the college experience blocks of variables as this is

sequentially the next set of blocks based on chronological time. I loaded the post-attendance block last because these variables happen after the students' exit Commuter State. The blocks are loaded sequentially related to the timing of enrollment so that the findings can provide an understanding of real-time default mitigation. I discuss this in greater detail within the regression model discussion of the analytical approach portion of this chapter. To this end, I added the college experience blocks to test whether these in-school variables added further understanding of what predicts default, addressing research question three. In summation, the model in research question three is organized from the most commonly used metrics, and items institutions cannot control, the pre-college variables, to the variables an institution may be able to control, the college experience, to additional commonly used metrics, post-attendance variables.

Pre-College Student Characteristics Block

The block of pre-college variables is typically used in the default literature (Gross et al., 2009). These factors are those the student brings with them to college; they are baseline conditions unaffected by the students' enrollment experience at the institution. The background characteristics include demographic information, SES factors, life situation (such as dependent or whether the students' have dependents), and incoming academic achievement (such as high school GPA and standardized test scores). The demographic variables include sex, race/ethnicity, age, and first-generation status. These variables, which are studied in previous research, comprise the students' background characteristics and help understand how a students' background is associated with defaulting. Ethnicity is perhaps the most studied characteristic in the loan default literature (Gross et al., 2009). Since previous research finds students of color and, specifically, African American and Hispanic students (Knapp & Seaks, 1992; Podgursky et al., 2002; Steiner & Teszler, 2003; Wilms et al., 1987) were more likely to default compared to

71

their White peers, I structured the race/ethnicity variable with White as the reference variable, and African American and Hispanic as stand-alone categories. The remainder of the reported race or ethnicity groups were relatively small individually. As such, I decided to combine the remaining groups into the "other URM" category.

Several scholars have found that as a students' age of enrollment in postsecondary education increases, so does the likelihood of ever defaulting on their loans (Podgursky et al., 2002; Looney & Yannelis, 2015; Steiner & Teszler, 2005; Woo, 2002). The findings informed my decision to include age at the point of enrollment as a continuous variable in the model. Parental education is another significant indicator of student default (Volkwein et al., 1998; Volkwein & Szelest, 1995). Existing research suggests that students with parents with higher levels of education were less likely to default compared to their first-generation college peers. Thus, my model's reference group for the categorical first-generation status variable is the students who are not first-generation (i.e., the students who have at least one parent who received a bachelor's degree), which aligns with Commuter State's determination of the first-generation status. The variable is categorical and not binary because students at Commuter State can have an unknown first-generation status. These students are categorized as unknown because they reported they did not know their parents' education status or did not report their parents' education status. In either case, these individuals comprise the group of unknown first-generation status students.

Parental income is included in the pre-college student characteristics block because the literature regarding family income and student default suggests the greater the socio-economic status of the students' family pre-college, the less likely a student will default (Hylands, 2014; Knapp & Seaks, 1992; Looney & Yannelis, 2015; Mezza & Sommer, 2015; Wilms et al., 1987;

72

Woo, 2002). These findings are incorporated into the model by including expected family contribution (EFC), Pell-eligibility, and total loan amount into my analyses. EFC is the Department of Education's calculated determination of what a family can afford to pay for college. EFC is a continuously scaled variable (to the dollar). I have scaled the variable to the nearest $1000 for interpretive purposes. For example, it is more meaningful and better associated with annual household incomes to say "with each $1000 increase in ability to pay" compared to "with each additional $1 of ability to pay." The Pell-eligible variable is a binary non-Pell and Pell categorical variable, which is determined by the Department of Education. University grant dollars are measured to understand the contribution they make to students' likelihood to default. The variable is formatted similarly to the EFC variable, starting with $0 and increasing by $1000 increments.

The life situation variables included in the pre-college characteristics block are dependency status and whether the student has dependents. The dependency status of a borrower helps understand the students' financial obligations beyond paying for college. This is a binary variable with the reference group as no dependents. Another way to understand a students' financial obligation is by measuring their dependency status. Dependent students, for this binary variable, are the reference group.

The final group of variables in this block includes the incoming academic achievement variables. This set of variables helps provide context to the academic readiness of these students before entering college. Research finds as high school rank, GPA, and standardized test scores increase, the likelihood of defaulting decreases (Podgursky et al., 2002; Steiner & Teszler, 2003; Woo, 2002). Commuter State does not collect high school rank for its students; however, it does collect GPA and standardized test scores. The incoming GPA (whether it is high school GPA for

73

first-year students or previous institution's GPA for new transfers) is rounded to increments of 0.1 for ease of interpretation. ACT composite test score is a continuously scaled whole number variable that is included in the model and did not need transformation.

College Experience Block

As the name infers, the college experience block contains the college experience variables. These variables are incorporated into the model to understand if and how these variables help explain student default after accounting for the variables in the pre-college block. It is important to note that the college experience variables are ordered by student term enrolled, such as first term, second term, and so on. However, it is not always the case that term 2 was immediately after term 1. For instance, a student could stop out for one semester (or more). After the student takes the semester off, they re-enroll and continue their enrollment the following five semesters, ultimately graduating. In this case, the student graduated in six semesters; however, it was over a seven semester span because of the one semester the student stopped out. From my data set I cannot discern whether students had gaps in enrollment. There are important implications to this data structure related to the interpretation of the findings. I discuss these implications in my discussion of regression model two.

The factors included in the college experience blocks are term GPA, term attempted and completed credits, and major. I rounded the continuously scaled variable term GPA to the nearest 0.1 increment. The GPA of a student helps to understand the students' academic success at a particular time during their enrollment history. The GPA provides insight towards student persistence and also towards their experience as a student, both of which are important factors to measure in the model. Credits attempted and earned, by term, are other measures of persistence. These scaled integer variables help provide another context for students' persistence and their

74

experience. A student's major is another way to measure their experience while enrolled. Volkwein et al. (1995) found that science or technology majors were incrementally less likely to default compared to their peers from other majors. Existing research informed the structure of the major variable such that non-STEM majors are the reference group for the categorical binary variable.

Post-Attendance Block

In the previous sections I discussed the pre-college and college experience blocks. In this section I discuss the third block of variables, the post-attendance block. This block includes two variables, graduation and total federal student loan amount. Similar to the dependent variable, default, the graduation variable is a binary categorical variable with an outcome of graduated or not graduated. This variable is important to include in the study due to its strong association with defaulting (Dynarski, 1994; Knapp & Seaks, 1992; Looney & Yannelis, 2015; Mezza & Sommer, 2015; Volkwein et al., 1998; Woo, 2002). The cumulative federal loan amount is a continuously scaled variable that I transformed to a scale of $1,000 increments, which replicates the structure of other recent research (e.g., Hillman, 2014; Mezza & Sommer, 2015). Recent studies concluded that student loan balances are generally not a significant predictor of student loan delinquency. However, where loan balance does make a difference there is evidence that student debt burden has an inverse relationship to the likelihood of defaulting (Dynarski & Kreisman, 2013; Hyland, 2014; Mezza & Sommer, 2015). The lower the debt amount, the higher the likelihood to default. Given this information, it is still vital and relevant to include loan debt levels within the model as it should be accounted for and mirrors the variables included in previous studies.

It is important to note that many of the variables in the study are likely inter-correlated. For example, family income is likely related to the amount borrowed. This inter-correlation can affect the stability of individual parameter estimates; however, I am more interested in how various blocks of variables affect the overall model fit. These changes in overall fit informed me about how predictive the post-, pre-, or intra-college variables were towards default, and suggest the degree to which institutional action can affect the ultimate outcomes of student default.

Analytical Approach

In the following section, I discuss the multifaceted data analysis process I used in this dissertation. The analytical approach began with a descriptive analysis of the variables. The descriptive analysis helped answer research question one: what are the descriptive characteristics of students at a nonresidential campus who do and do not default on their loans? The next section discusses the regression models, which addresses research questions two and three: what predicts default among students at Commuter State, and do institution-specific, student-level measures improve estimates of default probabilities for students at Commuter State?

Descriptive Analysis

At the outset, I conducted a descriptive analysis of the variables on all the data pooled across the years. The descriptive analysis compared defaulters to non-defaulters for all the variables in the analytic blocks described in the variables section. The descriptive analysis is important for two reasons. First, it provides information to support the regression, such as inferences about the variables' central tendencies and dispersion within the data set. The analysis provides an understanding of the ranges, averages, and outliers of the various variables in the model. It provides an understanding of the distribution of the data by outputs of simple counts for categorical variables, measures of shape for continuous variables, yields an understanding of

the scope of missing data, and provides the necessary metrics for calculating the uncontrolled

odds ratios for the categorical variables. This is an important aspect of the analysis for the

categorical variables because once I completed the regression analyses, I compared if the odds

ratios differed between the descriptive statistics and the regression models.

Second, the findings from the descriptive analysis provide insights into whether there is

anything substantively interesting about the population of defaulting students compared to the

non-defaulters, answering research question one. This is a key element because Commuter State

knows very little about the population of students who have defaulted (and, just as important,

those that have not defaulted). So any information about this population contributes to the overall

understanding of commuter student default at Commuter State. In addition, it provides a starting

point for the regression model by identifying particular variables that may have some predictive

power to identify defaulters. For example, one can easily compare the odds ratios of the

categorical variables from the descriptive analysis to the regression analyses for the same

categorical variables, which may yield different findings for the same data.

The Regression Model

Once I described the data set variables, my analysis moved to the second phase, the

logistic regression models. My study included three regression models. The first regression

model addressed research question 2 (What predicts default among students at Commuter

State?). Regression model B1 and model two address research question 3 (Do institution-

specific, student-level measures improve estimates of default for students at Commuter State?)

using different lenses. The purpose of regression model B1 was for a robustness check. I discuss

regression model B1 and my findings for the model in Appendix B. In the next section, I outline

the specific approach to model building utilized for my study. After I establish an understanding

of the technique and a justification for the block order, I then discuss the two different regression models.

Model Building

The regression approach utilized for my study is logistic regression. This approach is designed to conduct multiple regression models for binary outcome variables (Lomax & Hahs-Vaughn, 2012). My model's dependent variable (default) is binary and categorized as defaulted (value=1) or not defaulted (value=0). Applying ordinary least squares (OLS) to a binary outcome creates problems because OLS estimates are based on linear relationships between the independent and dependent variables (Lomax & Hahs-Vaughn, 2012). Since default is binary, it is problematic to apply a regression technique that assumes linear relationships to dichotomous outcomes. Similar to OLS, binary logistic regression provides outcome estimates per a set of given inputs; however, the output of the regression predicts the probably that the dependent variable will result in one of two outcomes, default and not defaulting (Lomax & Hahs-Vaughn, 2012).

In addition to utilizing a binary logistic approach, my research utilized a specific model building approach called block regression. Block regressions allow us to explore how each block of variables relates to the full model (Naes et al., 2013). The creation of the predictor blocks is based on the common theoretical ground resulting from careful consideration of the available research (Lomax & Hahs-Vaughn, 2012). The sequence of the blocks loaded into the model are specified *a priori*, again based on existing research (Lomax & Hahs-Vaughn, 2012). Specifically, I created the blocks in this study to address the second and third research questions.

The existing research findings discussed in chapter 2 and the previous variables section informed which variables I included in my study's blocks. In the following discussion of the two

regression models, I outline the order the blocks were loaded into the models. As discussed earlier, the predictor variables cluster into three groups: pre-college student characteristics, college experiences by semester enrolled, and outcomes. The first block included background characteristic variables. The variables I included in this block are the key characteristics existing research found associated with defaulting. I included the next set of blocks, college experiences by semester enrolled variables because some research suggests their connectedness to defaulting (Harrast, 2004; Podgursky et al., 2002; Steiner & Teszler, 2005). I included these blocks because they are the quantifiable, available institutional data from Commuter State. The final block, post-attendance, comprised the graduation variable and total loan amount from Commuter State. The block captured the contributions graduation and loan debt make to understanding default. The inclusion of graduation in the model aligns with the existing research methodology and is important because the research finds that graduation is one of the greatest predictors of default.

After I created the blocks of variables, I determined the loading sequence of the blocks into the models. I conditioned the loading sequence of the blocks into the regression models based on which of my research questions I sought to answer. In the following sections, I outline the two different regression models.

Regression Model One

Regression model one was structured to answer research question two: what predicts default among students at Commuter State? To address this question, I configured regression model one as close as possible to the national models because it informs whether the traditional predictors of defaulting are predictive for only Commuter State students. To this end, regression model one utilized the pre-college block variables and the post-attendance block, which are the most commonly used variables in national studies on the issue and those most readily available

79

in national datasets. The *a priori* sequence of the models was the pre-college student

characteristics block followed by the outcome block. The pre-college characteristics block is the

first block loaded into the model. The decision to load this block of variables first was to account

for the most important predictors of defaulting: the students' incoming characteristics. Just as

important, this is the typical approach of the existing national studies (Gross et al., 2009).

Loading the pre-college variables into the model in the first block provided an understanding of

how influential the set of variables are to defaulting absent the other variables in the model. After

the model measures the association between pre-college characteristics and default, I loaded the

post-attendance block of variables into the model. The decision to load the outcome variables

into the model accounts for one of the most important predictors of defaulting, whether a student

graduated or not. After accounting for pre-college characteristics, loading graduation into the

model provided an understanding of how influential graduation and debt are to defaulting and

structured the regression model similar to national studies. As I outlined in chapter 2, commuter

students are different from residential students. The primary purpose of regression model one

was to understand which traditional default factors are predictors of default among students at

Commuter State. In the next section, I discuss regression model two. Please refer to Appendix B

for my discussion of regression model B1.

Regression Model Two

Regression model two utilized all three blocks of data but in a slightly different

configuration. The *a priori* sequence of the model was similar to regression model one in that the

pre-college student characteristics block was added first. However, instead of the post-attendance

block following as in more traditional models of default, I now have access to more detailed

local student attendance data which I included next. Finally, I included the post-attendance block

of variables which again are consistent with more traditional studies of default. Regression model two is essentially identical to model one, but with the inclusion of college experience measures, thus adding local data into the mix, possibly expanding beyond the national models' ability to predict default.

For regression model two, the pre-college student characteristics set of predictors was the first block loaded into the model because existing research suggests that student background characteristics are a strong predictor of default, and it is the first data that colleges and universities have on students which is an important consideration if institutions want to use these models to predict and address future defaults. My results from this order identify how strongly associated the pre-college variables are with defaulting. After fitting the association between pre-college characteristics and default, the next set of blocks of predictors loaded into the models are the factors measuring college experiences by semester enrolled. The hypotheses are that the contribution of students' background characteristics to their likelihood to default should be accounted for before analyzing the connectedness between institutional data, the second set of blocks, and default. The sequence of the blocks is also a function of the order that the data comes to the institution, which is important because then my results provide an ability to measure default as the students' progress through their enrollment at Commuter State. The college experiences by semester enrolled blocks are loaded by order of chronological enrollment. This process helps me investigate research question three: do institution-specific, student-level measures improve estimates of default for students at Commuter State?

The final block of variables loaded into regression model two are the post-attendance variables. The decision to load the post-attendance variables was to account for one of the most important predictors of defaulting, whether a student graduated or not. Loading graduation into

the model as the last block explained how influential graduation and debt were to defaulting after accounting for the other variables in the model. Please refer to Appendix B for a robustness check of my model.

I outlined in the previous section my regression model two's approach to loading data into the model. For regression model two, I had to conduct a transformation to the variables included in the student experience blocks of the model to ensure the model included all the students for every term. It is typical to see students drop out semester over semester. For those students who dropped out, their term data is blank. For instance, every student in the data set was enrolled in term 1. Beginning with term 2, students started dropping out; thus, these students' data for term 2 is missing, so the regression model's N decreases to only those students still enrolled, those students who have term 2 enrollment data. I treated this as a missing data problem. To ensure all the students were included in the model, I utilized the dummy variable adjustment method by assigning zeros to those students who had missing data for term GPA, credits attempted, and credits earned, per term, and then coded a dummy enrollment flag, per term, to indicate the enrolled or not enrolled by term (Allison, 2002). Allison (2002) describes that the dummy variable adjustment method is not appropriate for actually missing data; however, is appropriate when missing data, unobserved values, do not exist. My approach, via the dummy variable adjustment method, to the missing data allowed every student to remain in each model iteration. In other words, I structured the model to calculate the associations of default for all the available data from Commuter State for every term added to the model. As such, regression model two provides information about whether dropping out (or persisting) is related to defaulting for every term in the model. Therefore, as I added additional terms to the model, the number of students remained constant. Regression model two tells us which students

should receive default interventions regardless of whether they are enrolled or not. It provides earlier information about default leavers prior to adding post-attendance data.

Overall, the purpose of the analytic modeling is twofold. One, it reveals how the predictors of default at Commuter State relate to what is already known from national studies. Two, it determines if institution-specific, student-level measures improve estimates of default probabilities for commuter students at a nonresidential campus. More specifically, a model that can predict default with institution-specific, student-level data suggests the addition of these factors to the model is useful for estimating default. In this situation, Commuter State could use this information to develop policies and programs to minimize default for future students. If the model is not able to accurately predict default with the addition of the college experience by semester enrolled blocks of variables, my findings suggest the institutional data does not improve estimates of default for commuter students and, therefore, the factors related to default are predicted by pre-college characteristics, graduating, other unmeasured factors, or a combination of these factors.

Below I provide the equation for the regression model two. The Y is the dependent variable, default. The blocks pre-college characteristics, college experiences, and post-attendance are the vectors of predictor variables within each block, and the bs are the coefficients in the equation. There are coefficients in the equation with more particular interest than others to address specific research questions. For example, the coefficients of the pre-college student characteristics block (b_s) is important to understand how the traditional predictors of default for Commuter State compare to results from previous studies. The subsequent blocks after the pre-college block include the college experience blocks and the post-attendance variables. The coefficients estimated for these blocks are important individually; however, the entire block is

even more important. For instance, if the model's overall explanatory power increases substantially after adding a specific set of college experience variables, an important point during the borrowers' enrollment period is identified. This is an opportunity to dig into the individual factors within blocks to understand what enrollment factors are strongly associated with default, helping to understand whether institution-specific, student-level measures improve estimates of default probabilities for commuter students and which variables within the block are, particularly, important.

$$\Pr[\text{Default}=1] = 1/[1+\exp(b_o + \mathbf{b}_s\text{Pre-College Characteristics} + \mathbf{c}_s\text{College Experiences} + \mathbf{d}_s\text{Post-Attendance})]$$

Default = dependent variable

b_o = intercept

$\mathbf{b}_s$ = vector of coefficients on Pre-College

Characteristics

$\mathbf{c}_s$ = vector of coefficients on College Experiences

$\mathbf{d}_s$ = vector of coefficients on Post-Attendance

The block regression analysis results for the population provide an understanding of the impact of the overall model and, specifically, the coefficients of interest. The blocks are loaded into the model, beginning with pre-college student characteristics followed by the data for the students' experiences while enrolled at the institution. The final block loaded is the post-attendance variables. With each additional block, I observed how the coefficients on the key predictors and the overall model fit changed. This provides information about which blocks matter to the overall fit of the model and the degree to which variables within the blocks are correlated with one another.

Limitations

As with any study, this one also has several limitations. First, I do not have access to robust outcome data such as employment data, income data (individual and household income data), other debt data beyond student loans, or the borrowers' life circumstances (being married, having dependents). These types of data more fully reflect the borrowers' financial situation, or ability-to-pay, after exiting higher education. Without this data, the model cannot account for these factors' association with defaulting.

An additional limitation is that the study does not capture borrowers' complete debt situation. Debt obligations such as private and personal student loans above and beyond their federal loans are not included. Nor does it capture the borrowers' other categories of debt, such as auto loans, mortgages, and credit card debt. The study should include all of the borrowers' debt obligations to understand the complete picture of a borrowers' ability to repay their federal student loans successfully. For example, if borrowers' earned income is less than their entire repayment obligation, then they are required to make decisions on which debts they will repay and which will go into default. These choices could associate with a borrowers' current life situation more than their background or their collegiate experience. Another limitation of the study is its inability to measure the borrowers' perspective on repayment. Individuals may default on their debt obligations because they cannot afford to make the payments; however, in other situations, borrowers may choose not to repay their loans even if they do have the financial means to repay successfully.

Beyond limitations in the availability of additional data, there are limitations to the measurement of default and the number of defaulters in the data set. Measuring default just three years after exiting higher education is a relatively short window. How would the results differ if

the Department of Education extended the period? Would they differ? The number of defaulters in the data set is relatively low. If the period extended to five years, would there be more defaulters? Also, the low number of defaulters could be sensitive to individual outcomes as opposed to overall trends. If there were more defaulters in the data, the data might better represent trends that emerge that are associated with default.

Finally, there are limitations on correlations between predictors. Specifically, a problem of endogeneity is present with specific input measures directly linked to graduation, such as credits completed and college GPAs. The problem of endogeneity means that the specific estimates of the independent variables may be adversely affected. Although the contribution of specific variables to defaulting is affected, the ability to calculate probabilities of default remains strong by the model design and, overall, this aspect of the study is more important than understanding the unique associations of individual factors and default.

In this chapter, I outlined the key components of the methodology of my research project, including a discussion of the population and sampling procedure, the data sources, the variables, analytic methods, and the key limitations of the study. In chapter 1, I outlined the significance of the study, the contribution of the study, and briefly introduced the existing findings of research related to the study. In chapter 2, I provided research evidence to support many of the decisions I made for the methodological approach and the design of my study. In the next chapters, I discuss the research findings, the implication of the findings, and the ways this study can be improved with future research.

CHAPTER 4: RESULTS

The purpose of my quantitative case study is to understand the predictors of student loan default at a nonresidential campus and, going a step further, if institutional-specific data can improve estimates of default probabilities. In this chapter, I present my study's findings organized by my research questions:

1) What are the characteristics of Commuter State students who do and do not default?

2) What predicts default among students at Commuter State?

3) Do institution-specific, student-level measures improve estimates of default for students at Commuter State?

Addressing research question one, in the first section of this chapter, I discuss the characteristics of defaulted students from Commuter State. The subsequent section compares the results for this study to the traditional predictors of default from previous studies. This discussion addresses research question two. The final section of this chapter, highlighting question three, discusses the findings from the block logistic regression model, which seeks to understand if institution-specific, student-level measures improve estimates of default probabilities. Before I delve into the research questions, I describe the aggregate characteristics of the Commuter State students included in this study. It is important to set the context of the overall population of the students within the study before delving into research question one which discusses the characteristics of the defaulted population.

Characteristics of Students at Commuter State

My study's population included all nine cohorts of 3-year default data for Commuter State, spanning from the 2009 cohort to the 2017 cohort. The entire data set included 13,181 borrowers. In the following sections, I discuss the study's population—students at Commuter

State who borrowed federal student loans to cover college costs—by the category of variables outlined in chapter 3: pre-college characteristics, college experience by semester enrolled variables, and post-attendance variables.

The following discussion provides insight to the population's aggregate pre-college student characteristics (see Table 3 and Table 4). To ensure clarity, I explain the tables and then discuss the results provided in the tables. Table 3 includes the descriptive analyses for the continuous variables in my study. Table 4 provides the descriptive analyses, odds, and odds ratios for the categorical variables in my study. It is important to note Table 4 is structured to present the categorical variables in odds ratios to align with the odds ratio outputs for the regression models. I calculated the odds column in Table 4 for each category of each variable. I calculated the odds by dividing the number defaulted in each category by the number not defaulted. I then calculated the odds ratios for each category by dividing the category's odds by the reference category's odds.

Table 3. *Descriptive Statistics of Continuous Variables for Defaulted and Not Defaulted Borrowers*

	Total				Not Defaulted				Defaulted			
	N	Mean	Median	Std. Dev.	N	Mean	Median	Std. Dev.	N	Mean	Median	Std. Dev.
Post-Attendance												
Total Loan Amount	13181	$16,947	$13,080	$14,453	12426	$16,986	$13,080	$14,437	755	$16,302	$10,791	$14,703
Pre-College Student Characteristics												
EFC	13116	$7,620	$3,240	$14,263	12364	$7,855	$3,240	$14,530	752	$3,766	$1	$7,826
Commuter State Grant	13181	$3,364	$2,570	$3,373	12426	$3,340	$2,570	$3,387	755	$3,774	$3,038	$3,120
Incoming GPA	13181	3.3	3.3	0.4	12426	3.3	3.3	0.4	755	3.2	3.2	0.4
ACT Composite Test Score	13181	21.3	20.0	3.0	12426	21.3	20.0	3.0	755	20.7	20.0	2.5
Age	13181	23.0	20.0	7.3	12426	22.9	20.0	7.2	755	25.2	22.0	8.9
College Experiences by Semester Enrolled												
Term 1 GPA	13175	2.8	3.0	1.0	12420	2.8	3.0	1.0	755	2.3	2.6	1.2
Term 1 Hours Attempted	13175	11.2	12.0	3.2	12420	11.2	12.0	3.2	755	10.7	12.0	3.4
Term 1 Hours Earned	13175	9.1	10.0	4.2	12420	9.2	10.0	4.2	755	7.5	9.0	4.4
Term 2 GPA	12416	2.7	2.9	1.1	11740	2.7	2.9	1.1	676	2.2	2.5	1.2
Term 2 Hours Attempted	12416	11.1	12.0	3.5	11740	11.1	12.0	3.5	676	10.7	12.0	3.3
Term 2 Hours Earned	12416	9.0	10.0	4.5	11740	9.1	10.0	4.4	676	7.4	8.0	4.6
Term 3 GPA	11275	2.7	3.0	1.1	10709	2.7	3.0	1.1	566	2.3	2.6	1.2
Term 3 Hours Attempted	11275	9.6	10.0	4.1	10709	9.6	10.0	4.1	566	9.6	10.0	3.7
Term 3 Hours Earned	11275	8.0	8.0	4.6	10709	8.1	8.0	4.6	566	6.7	6.0	4.4
Term 4 GPA	10548	2.8	3.0	1.0	10060	2.8	3.0	1.0	488	2.2	2.6	1.2
Term 4 Hours Attempted	10548	11.1	12.0	3.6	10060	11.1	12.0	3.6	488	10.6	12.0	3.7
Term 4 Hours Earned	10548	9.6	10.0	4.5	10060	9.7	10.0	4.4	488	7.4	7.0	4.8
Term 5 GPA	9870	2.8	3.0	1.0	9452	2.8	3.0	1.0	418	2.4	2.6	1.2
Term 5 Hours Attempted	9870	10.3	12.0	4.1	9452	10.3	12.0	4.1	418	10.4	12.0	3.7
Term 5 Hours Earned	9870	9.0	9.0	4.6	9452	9.0	9.0	4.6	418	7.7	8.0	4.5
Term 6 GPA	9162	2.8	3.1	1.0	8795	2.9	3.1	1.0	367	2.4	2.7	1.2
Term 6 Hours Attempted	9162	9.9	11.0	4.0	8795	10.0	11.0	4.0	367	9.8	10.0	3.7
Term 6 Hours Earned	9162	8.8	9.0	4.5	8795	8.8	9.0	4.5	367	7.4	7.0	4.4

Table 4. *Descriptive Statistics of Categorical Variables for Defaulted and Not Defaulted Borrowers*

		Total	Not Defaulted		Defaulted		Odds	Odds Ratio
		N	N	%	N	%		
Post-Attendance								
Degree Status		13181	12426		755			
	Graduated	7867	7686	97.7%	181	2.3%	0.024	
	No Degree from Subject Institution	5314	4740	89.2%	574	10.8%	0.121	5.1
Pre-College Student Characteristics								
Sex		13179	12424		755			
	Female	7315	6912	94.5%	403	5.5%	0.058	
	Male	5864	5512	94.0%	352	6.0%	0.064	1.1
Race/Ethnicity		13181	12426		755			
	White	8067	7685	95.3%	382	4.7%	0.050	
	Black	2244	2040	90.9%	204	9.1%	0.100	2.0
	Hispanic	448	419	93.5%	29	6.5%	0.069	1.4
	Other URM	2216	2093	94.4%	123	5.6%	0.059	1.2
	Unknown	206	189	91.7%	17	8.3%	0.090	1.8
Pell-Eligible		13181	12426		755			
	Non-Pell	6020	5807	96.5%	213	3.5%	0.037	
	Pell	7161	6619	92.4%	542	7.6%	0.082	2.2
First-Generation Status		13181	12426		755			
	Non-First-Generation	6907	6578	95.2%	329	4.8%	0.050	
	First-Generation	5596	5228	93.4%	368	6.6%	0.070	1.4
	Unknown	678	620	91.4%	58	8.6%	0.094	1.9
Dependency Status		13147	12394		753			
	Dependent	8518	8137	95.5%	381	4.5%	0.047	
	Independent	4629	4257	92.0%	372	8.0%	0.087	1.9
Student has Dependents		13181	12426		755			
	No Dependents	12803	12074	94.3%	729	5.7%	0.060	
	Has Dependents	378	352	93.1%	26	6.9%	0.074	1.2
College Experiences by Semester Enrolled								
Major		13181	12426		755			
	Non-STEM Major	9121	8557	93.8%	564	6.2%	0.066	
	STEM Major	4060	3869	95.3%	191	4.7%	0.049	0.7

The average incoming age of the borrowers is 23, with a median age of 20. The average expected family contribution (EFC) , which is the Department of Education's calculated determination of what a family can afford to pay for college, is $7,620, however. In their first year, the students received an average of $3,364 in grant aid from Commuter State. This is financial aid dollars received from the institution that does not need to be repaid. The students' average incoming GPA—the admitted GPA on their admissions application—is 3.3. Their average ACT composite test score is 21.3. As indicated on Table 4, 56% self-reported as females.[9] The population self-reported as predominantly White (61%), followed by Black (17%), other URM (17%), Hispanic (3%), and unknown (2%).[10] The majority of students are Pell-eligible (54%), which means they qualified via their FAFSA results for the need-based federally provided Pell Grant. More than 50% of the population indicated they were non-first-generation students (52%), while 5% had an unknown first-generation status which meant these students did not provide this data to Commuter State. Almost two-thirds (65%) of the population indicated being a dependent, meaning they are not financially independent, and 12,803 (97%) of the population had no dependents at the time of their initial enrollment at Commuter State.

Data from Commuter State also provides insights into the average college experience metrics for the Commuter State population. Over two-thirds (69%) of the students are non-STEM majors. The average student has a first term GPA of approximately a 2.8, with about 11 hours attempted and 9 hours earned. Of the students who returned for their second semester, their average GPA and credits attempted and earned almost mirrored that of the first semester. In term

[9] The only options for sex in the data set were female and male which is important to note because populations are no longer binary.
[10] As outlined in chapter 3, the Other URM category is an aggregate of the remainder of the race and ethnicity groups beyond Black and Hispanic and the category unknown includes students who did not report their race/ethnicity.

3, average term GPA is 2.7; however, average credits attempted decreased to 9.6 with an average credits earned of 8. The remainder of the semesters are displayed in Table 3.

The final category of variables is post-attendance. The first of the two variables in this category is graduation. The majority of students in this population graduated from Commuter State (60%). The second variable is average federal loan amount, which is the aggregate amount of federal loans each student owed at the time when their repayment process began. The population's average federal loan amount is $16,947. With this aggregate understanding of the population, in the next section I discuss the characteristics of students from Commuter State who defaulted.

Characteristics of Commuter State Students who do and do not Default

To address research question one, I calculated descriptive statistics to understand the characteristics of students from a nonresidential campus who do and do not default. The descriptive analyses included all three categories of variables: pre-college student characteristics, college experiences by semester enrolled, and post-attendance. In the following section, I discuss my overall findings for each category of variables.

With regard to pre-college student characteristics (i.e., sex, race/ethnicity, age, EFC, Pell-eligibility, Commuter State Grant award, first-generation status, independent status, student has dependents, incoming GPA, and ACT composite score), defaulters have more need, on average, than their not defaulting peers. Further, Commuter State defaulters are more likely to be Pell-eligible, Black, have an unknown first-generation status, and be independent. With regard to college experiences by semester enrolled (i.e., STEM major, term GPA, term hours attempted, and term hours earned), my results show differences in term GPAs and hours earned per semester between defaulters and those not defaulting. Defaulters, on average, have lower term

92

GPAs and, for almost every semester, fewer hours attempted and earned. In addition, defaulters are more likely to study non-STEM majors than their peers in good repayment status. Finally, for post-attendance variables (i.e., graduation status and total federal student loan amount), defaulters borrowed less and are less likely to receive a degree from Commuter State than those not defaulting. It is important to note, for research question one, my findings for each factor are mutually exclusive of the other factors. The regression model discussed later in this chapter provides insights into the factors associated with defaulting when accounting for all of the other factors in the model. The following three sections, organized by the three blocks of variables, present my findings from my descriptive analysis in greater detail.

Pre-College Student Characteristics

In this section, I discuss pre-college characteristics of defaulted students from Commuter State. Through descriptive analyses, I show the important pre-college student characteristics of defaulters from Commuter State are EFC, Pell-eligibility, race/ethnicity, first-generation status, and dependency status. Previous research finds these pre-college characteristics are important factors of default (Dynarski, 1994; Flint, 1997; Hillman, 2014; Knapp & Leaks, 1992; Looney & Yannelis, 2015; Monteverde, 2000; Podgursky, et al., 2002; Scott-Clayton, 2018; Steiner & Teszler, 2003, 2005; Volkwein & Szelest, 1995; Volkwein & Cabrera, 1998; Wilms et al., 1987; Woo, 2002). In the following sections, I discuss the findings for each of the above pre-college student characteristics for Commuter State.

EFC (Expected Family Contribution)

Defaulted borrowers' EFC (M = $3,766) is, on average, half of their not defaulting peers' EFC (M = $7,855). Of the defaulters, approximately half have an EFC of $0, the lowest EFC from the FAFSA calculations (see Table 1). The literature regarding family income and student

default suggests, pre-college, the higher the family income, the greater the socioeconomic status of the student's family, the less likely it is a student will default (Hylands, 2014; Knapp & Seaks, 1992; Looney & Yannelis, 2015; Mezza & Sommer, 2015; Wilms et al., 1987; Woo, 2002). My findings, then, are consistent with previous research. At Commuter State, family income matters in that defaulted students, on average, have more need than those not defaulting.

College Experiences by Semester Enrolled

In this section, I discuss the findings for the characteristics of Commuter State defaulters for the college experiences by semester enrolled variables. My results indicate there are differences between defaulters and those not defaulting in term GPAs and hours earned per semester. Further, whether a student is a STEM major is another important college experience default characteristic. It is important to note that this category of variables, college experiences by semester enrolled, are not traditional variables included in previous studies. These institution-specific variables are local variables; they are not available in administrative national data sets and, therefore, are not included in national default studies.

Term GPAs

On average, defaulted students have approximately a half-point lower term GPA(see Table 1). This trend is across all terms, give or take a 0.1 point of a term GPA. The greatest difference between average term GPAs for defaulters and those not defaulting is in term 4, with a 0.60 difference. In contrast, the smallest difference is for term 5 at 0.45. A half-point difference in GPA is quite large, considering GPA is on a four-point scale at Commuter State.

Major

The final college experience variable that is an important characteristic of students who defaulted is major. Over 6% of non-STEM majors defaulted compared to 4.7% of STEM majors. The odds of STEM majors defaulting is 30% less than non-STEM majors.

Post-Attendance

Both of the post-attendance variables, total loan amount and graduation, are essential characteristics associated with defaulting. Defaulters, on average, have lower total loan amounts. In other words, those who have a higher loan balance are less likely to default. The assumption is that students who persist to graduation enroll in more years of college than their not graduated peers, thus taking on more debt with the additional years of education. Graduated students are less likely to default than peers who did not earn a degree from Commuter State. In the following sections, I discuss these findings further.

Total Loan Amount

As described in Table 3, defaulters borrowed approximately $600 less, on average, than those not defaulting. Looking at another measure of central tendency, the median, further supports my finding. The median loan amount for defaulters ($10,791) is almost $2,300 less than the median of those not defaulting ($13,080). The mean and median for the students not defaulting are almost the same, suggesting several defaulters borrowed a lot of money. My distribution analysis results and median values provide evidence of the skewed distribution and suggest that very high total loan amounts may be masking the differences in total borrowed when looking only at means. Given this skewed distribution, the median may provide a better understanding of the difference in total borrowing between defaulters and those not defaulting.

Student debt is inversely related to the likelihood of defaulting (Dynarski & Kreisman, 2013; Hyland, 2014; Mezza & Sommer, 2015). The lower the debt amount, the greater the likelihood of default. The median difference for Commuter State defaulters and those not defaulting in total loan amount of almost $2,300, with defaulters having the lower median borrowed amount, aligns with previous research indicating that the total loan amount is important and is inversely related to default. As I noted earlier, one explanation of defaulters having lower loan amounts than their not defaulting peers is the number of years enrolled (Dynarski & Kreisman, 2013; Hyland, 2014; Mezza & Sommer, 2015). For example, students who graduate will enroll in more years of higher education, thus taking more loan debt than their defaulted peers who have dropped out and, as a result, enrolled in fewer years of higher education resulting in taking on less debt.

The borrowers' wealth may affect their ability to repay their debt. For defaulters, the median EFC is $1, meaning almost half of the defaulted population has the lowest possible EFC. These low EFC students receive Pell grants and institutional need aid, thus lowering their out-of-pocket costs and resulting in a lower amount of student loan aid needed to cover their enrollment costs. My results suggest that it could be more difficult for the less wealthy group to repay their debt even with a lower debt load.

Degree Status

Students who did not receive a degree from Commuter State are more likely to default than their graduated peers. As indicated in Table 4, 10.8% of students in the category of *no degree from subject institution* defaulted compared to just 2.3% for graduates. The greatest odds ratio relative to the reference group is for the graduated categorical variable. Borrowers with no

degree from Commuter State's odds of defaulting is 5.1 times higher than their graduated peers. The next largest odds ratio is 2.2, which is for the Pell-eligible variable.

The single strongest predictor of not defaulting is postsecondary degree completion (Dynarski, 1994; Knapp & Seaks, 1992; Looney & Yannelis, 2015; Mezza & Sommer, 2015; Volkwein et al., 1998; Woo, 2002). Researchers' findings consistently conclude that graduating is strongly linked to not defaulting. My study's findings are consistent with this previous research. As noted, graduation has the largest odds ratio of all the categorical variables, confirming graduation is the single strongest default predictor at Commuter State

In this section, I discussed my findings related to research question one addressing the characteristics of commuter students at a nonresidential campus who do and do not default on their loans. My results suggest that pre-college, during college, and post-college characteristics are important factors to understand Commuter State student default. In this section, I discussed my descriptive analysis results. My results provide information about the key characteristics of students who defaulted.

My results highlight the additional benefits that emerge from studying default locally, as a case study that may get washed out within the national context. Specifically, by studying one institution's local data, I was able to study the differences for those who do and do not default for college experience variables. However, it is important to note, since research question one focused on descriptive analyses, my findings for each variable are unique and not connected to the other variables' findings within my model. Regression modeling is a technique to help us understand the significant predictors of defaulting when controlling for the other variables in the model. In the following section, I discuss the results of regression model one, which includes the traditional characteristics of defaulting.

What Predicts Default among Students at Commuter State?

In this section, I address research question two:

What predicts default among students at Commuter State?

To address this question, I utilized largely similar variables as national studies. The regression model included the pre-college student characteristics block of variables and the post-attendance block of variables. The *a priori* sequence of the models was the pre-college student characteristics block followed by the outcome block. This is regression model one that I outlined in chapter 3. The primary purpose of regression model one was to understand which variables were predictors of default at Commuter State by including the traditional default model variables that are included in national studies. The results help answer which factors are important to Commuter State defaulters. My regression model one results find that the statistically significant predictors of defaulting are sex, EFC, the amount of grant aid from Commuter State, first-generation status, having no incoming admitted GPA, the total loan amount, and graduating. Race/ethnicity is perhaps the most studied characteristic in default research (Gross et al., 2009) and, overwhelmingly, it is a predictor of default. All other factors held constant, Commuter State differs from the national studies in that race/ethnicity is not a statistically significant predictor of default. In the following sections, I discuss the findings from regression model one, addressing what predicts default among students at Commuter State.

Pre-College Student Characteristics

Previous research finds pre-college characteristics are important factors of default (Dynarski, 1994; Flint, 1997; Hillman, 2014; Knapp & Leaks, 1992; Looney & Yannelis, 2015; Monteverde, 2000; Podgursky, et al., 2002; Scott-Clayton, 2018; Steiner & Teszler, 2003, 2005; Volkwein & Szelest, 1995; Volkwein & Cabrera, 1998; Wilms et al., 1987; Woo, 2002). My

regression model one results find that Commuter State defaulters' important pre-college student characteristics are sex, EFC, the amount of grant aid received from Commuter State, first-generation status, and whether a student had an incoming GPA from their previous institution. In the following sections, I first discuss in greater detail the results of my findings. I conclude the chapter by discussing the statistically significant factors with only the pre-college student characteristics loaded into the model but lost their statistical significance when I loaded the post-attendance block into the model. This discussion is important because it highlights the factors that the traditional default models find significant but not significant for Commuter State defaulters.

Sex

Sex is a statistically significant predictor of default at Commuter State when holding all else constant. The odds of males defaulting is 1.2 times greater than females. The variable is a statistically significant predictor in both the model with only pre-college characteristics and the entire model that included post-attendance variables (see Table 5).

Table 5. *Regression Model 1: The Traditional Default Model*

	Model 1.1: Pre-College Student Characteristics		Model 1.2: Post-Attendance	
Pre-College Student Characteristics				
Male	1.301	***	1.211	*
Race/Ethnicity (White)				
Black	1.497	***	1.130	
Hispanic	1.329		1.120	
Other URM	1.038		1.050	
Unknown	1.286		1.124	
Age	1.016	*	1.007	
Scaled EFC (1000s)	0.969	***	0.976	**
Pell-Eligible	1.304	*	1.223	
Scaled Commuter State Grant (1000s)	1.005		1.036	*
First-Generation Status (Non-First-Generation)				
First-Generation	1.239	**	1.201	*
Unknown	1.266		1.318	
Independent Student	1.156		1.044	
Student has Dependents	0.683		0.657	
Rounded Mean Admitted GPA (0.1s)	0.684	***	0.856	
Missing admitted GPA	1.115		1.411	***
Mean ACT Composite	0.990		0.969	
Missing ACT Composite Score	1.035		1.176	
Post-Attendance				
Not graduated from Commuter State			6.252	***
Total Loan Amount (1000s)			1.019	***
Intercept	0.112	***	0.025	***
N	13108		13108	

Overall Model Outputs	X^2	**-2 Log likelihood**	X^2	**-2 Log likelihood**
Omnibus Test of Model Coefficients	234	5525	622	5136
Hosmer & Lemeshow Test	4.814		2.251	
Cox & Snell R2	0.018		0.046	
Nagelkerke R2	0.050		0.130	

Note: Reference group or scaling within parentheses, significance level p<.001"***"; p<.01"**"; p<.05"*"

EFC (Expected Family Contribution)

EFC is a statistically significant predictor of defaulting for Commuter State students. Borrowers' who have $1000 more in EFC are 2.4% (p<.01) less likely to default. My results from regression model one show family income matters in that defaulted students, on average, had more need than their not defaulting peers. The higher the family income, the greater the socioeconomic status of the student's family, the less likely it is a student will default (Hylands, 2014; Knapp & Seaks, 1992; Looney & Yannelis, 2015; Mezza & Sommer, 2015; Wilms et al., 1987; Woo, 2002).

Commuter State Grant Aid

My results from regression model one indicate grant aid provided by Commuter State is a significant predictor of default, holding all else constant. It is an inverse relationship between grant aid and default. Borrowers who have $1,000 more in Commuter State grant aid are 3.6% (p<.05) more likely to default. This result could relate to EFC. As family need increases, measured by lower EFC, Commuter State awarded more institutional grant aid. However, the variable also includes merit aid as part of the aid awarded. Due to the grant aid variable's aggregated structure in the data set, I cannot differentiate the need aid from merit aid.

First-Generation Status

First-generation status is a statistically significant predictor of default for Commuter State students. The odds of defaulting for first-generation students is 1.2 times higher than non-first-generation students. This aligns with previous research findings, which finds parental education is a significant indicator of student default (Volkwein et al., 1998; Volkwein & Szelest, 1995).

Missing Admitted GPA

Not all students within the data set had an admitted GPA in their student records. To understand if this is a significant predictor of default, I created a binary dummy variable to measure the effect of having a GPA on file. Subsequently, the odds of students defaulting without an admitted GPA on file is 1.4 (p<.001) times greater than the students with a GPA. Scholars find that academic preparation—defined as high school rank, high school GPA, and standardized test scores—significantly contributes to loan default. As high school rank, GPA, and standardized test scores increase, the likelihood of defaulting decreases (Podgursky et al., 2002; Steiner & Teszler, 2003; Woo, 2002). This may be a finding unique to the data set at Commuter State. The students without admitted GPAs may apply to specific situations for students at Commuter State. This aspect of the data set provides opportunities for future avenues of exploration.

Important Traditional Default Model Pre-College Student Characteristics that were not Important at Commuter State

The pre-college student characteristics for Commuter State students including race/ethnicity, age, Pell-eligibility, and admitted GPA are statistically significant with only the pre-college student characteristics loaded into the regression model. These factors became non-significant when I loaded the post-attendance block.

The greatest difference emerged with race/ethnicity not being a statistically significant predictor of Commuter State borrowers defaulting. Initially, race/ethnicity is significant for only Black students. The odds of defaulting for Black students is 1.5 (p<.001) times greater than Whites. As shown in Table 5, once I added the post-attendance variables of graduation and total loan amount, the variable is no longer statistically significant.

The next largest pre-college predictor of defaulting is the admitted GPA. Each decrease of 1 point in admitted GPA increased the odds of defaulting by 32% (p<.001). Similar to the race/ethnicity variable, the addition of the post-attendance variables eliminated the predictive significance of admitted GPA.

Commuter State students' Pell-eligibility is initially a significant predictor of default. The odds of defaulting for Pell-eligible students is 1.3 (p<.05) times greater than non-Pell-eligible students. The Pell-eligible status became non-significant in the full regression model. Finally, age is statistically significant with the first pre-college block loaded into the model, but similar to the previous variables discussed, it becomes non-significant in the full model.

Post-Attendance

In this section, I discuss my regression results for the post-attendance variables. My results show that the two variables within the post-attendance category (i.e., total loan amount and graduation) are important characteristics of Commuter State defaulters. I discuss my findings in greater detail in the following sections.

Total Loan Amount

My results from regression model one find that total loan amount is a statistically significant predictor of default for Commuter State students. Borrowers' who have $1,000 more in total loan amount are 1.9% (p<.001) more likely to default, thus, a positive relationship between total loan amounts and defaulting.

Degree Status

The single strongest predictor of not defaulting is postsecondary degree completion (Dynarski, 1994; Knapp & Seaks, 1992; Looney & Yannelis, 2015; Mezza & Sommer, 2015; Volkwein et al., 1998; Woo, 2002). The odds of defaulting for students who did not receive a

degree from Commuter State are 6.3 times greater than those who graduated. This is the strongest predictor of not defaulting for Commuter State Students.

The previous section discussed the regression model one results for the traditional default model—the default model utilized by previous studies of national data sets—applied to my case study institution Commuter State. My results show which variables of the traditional model are statistically significant predictors for Commuter State defaulters and, just as important, which traditional default variables are not statistically significant at Commuter State. My results suggest the characteristics of commuter students who default are different from the findings for defaulters in national data sets and, thus, suggests local institutional data may add additional information to increase our ability to predict default for commuter students. In the following section, I investigate the results of adding institution-specific local data to the regression model.

Do Institution-Specific, Student-Level Measures Improve Estimates of Default Probabilities?

This section addresses my third research question:

Do institution-specific, student-level measures improve estimates of default for students at Commuter State?

Overall, the purpose of the analytic modeling for research question three is to determine if institution-specific, student-level measures improve estimates of default probabilities for commuter students at a nonresidential campus. My results show institution-specific, student-level measures improve estimates of defaulting. In particular, my results show that even after controlling for pre-college characteristics, student-level measures such as term GPAs, hours attempted, and hours earned can increase our understanding of defaulting. In the following sections, I discuss my findings for regression model two, and how they further help explain

commuter student default at a nonresidential campus. My results for regression model two confirm student-level measures can increase the variance explained sooner in the regression model for defaulting at Commuter State and, specifically, it provides predictors of default for enrolled Commuter State students, semester by semester. Further, my results from regression model two can also provide point-in-time default probabilities for Commuter State students who dropped out. The findings from regression model two identify which students should receive default interventions to mitigate defaults. My findings provide a framework other institutions could apply to both their currently enrolled and stopped out student populations to identify which students and when they are at risk of defaulting.

Regression Model Two

In this section, I discuss my findings for regression model two. As outlined in chapter three, regression model two keeps all the borrowers in the data set regardless of their term enrollment status. The model results indicate whether leaving at any point in time (measured by terms) during the students' enrollment can predict default probabilities and can target stopped-out students as well as those who remain enrolled. Not all students graduate from college and, therefore, stop out of higher education at various points. This is indicated by the gradual decrease of over 4,000 students from model B1.1 to model B1.8 in regression model B1 in Appendix B. To measure the impact of dropping out in addition to the college experience measures of term GPA, hours attempted, and hours earned, I created a non-enrollment variable (term X not enrolled) for terms 2 thru 6 for regression model two, represented in Table 6. This non-enrollment variable took on a value of 1 if students were not enrolled in that term and 0 if they were enrolled. My research findings from regression model two show that adding institution-specific, student-level measures improves estimates of default probabilities for my case study

institution, Commuter State. Thus, providing the opportunity to consider a framework for a default intervention model for students who are still enrolled and those who left Commuter State. In the following section, I discuss the results for regression model two.

Table 6. *Regression Model Two: Logistic Regression Estimates Pre-College Student Characteristics, College Experiences by Semester, and Post-Attendance (Expressed in Odds Ratios)*

	Model 2.1: Pre-College Student Characteristics		Model 2.2: College Experience Semester 1		Model 2.3: College Experience Semester 2		Model 2.4: College Experience Semester 3		Model 2.5: College Experience Semester 4		Model 2.6: College Experience Semester 5		Model 2.7: College Experience Semester 6		Model 2.8: Post-Attendance	
Pre-College Student Characteristics																
Male	1.301	***	1.289	**	1.262	**	1.254	**	1.245	**	1.233	**	1.226	*	1.197	*
Race/Ethnicity (White)		***														
Black	1.497	***	1.264	*	1.193		1.150		1.115		1.091		1.080		0.988	
Hispanic	1.329		1.229		1.211		1.153		1.148		1.146		1.131		1.059	
Other URM	1.038		1.016		1.009		0.990		0.993		0.998		0.992		1.033	
Unknown	1.286		1.231		1.186		1.095		1.127		1.100		1.107		1.124	
Age	1.016	*	1.017	*	1.017	*	1.017	*	1.018	**	1.017	*	1.018	**	1.011	
Scaled EFC (1000s)	0.969	***	0.973	***	0.973	***	0.973	***	0.973	***	0.974	***	0.974	***	0.976	**
Pell-Eligible	1.304	*	1.189		1.167		1.133		1.111		1.121		1.128		1.128	
Scaled Commuter State Grant (1000s)	1.005		1.028		1.028		1.030		1.031		1.031		1.031		1.035	*
First-Generation Status (Non-First-Generation)																
First-Generation	1.239	**	1.222	*	1.216	*	1.218	*	1.219	*	1.222	*	1.225	*	1.200	*
Unknown	1.266		1.246		1.239		1.242		1.232		1.250		1.278		1.295	
Independent Student	1.156		1.150		1.162		1.166		1.142		1.144		1.127		1.051	
Student has Dependents	0.683		0.681		0.683		0.687		0.667		0.671		0.672		0.674	
Rounded Mean Admitted GPA (0.1s)	0.684	***	0.823		0.896		0.936		0.990		1.004		1.017		1.057	
Missing admitted GPA	1.115		1.157		1.209	*	1.252	*	1.288	*	1.328	**	1.320	**	1.421	***
Mean ACT Composite	0.990		0.997		0.996		0.994		0.990		0.987		0.985		0.971	
Missing ACT Composite Score	1.035		1.088		1.128		1.135		1.162		1.161		1.161		1.226	

Table 6 (cont'd)

College Experiences by Semester Enrolled

	C1	C2	C3	C4	C5	C6	C7
STEM Major	0.802 *	0.776 **	0.753 **	0.743 **	0.737 **	0.736 **	0.745 **
Term 1 GPA (0.1s)	0.783 ***	0.862 **	0.871 **	0.879 *	0.885 *	0.894 *	0.893 *
Term 1 Hours Attempted	1.029	1.012	1.008	1.004	1.003	1.004	1.004
Term 1 Hours Earned	0.968 *	0.974	0.981	0.988	0.989	0.988	0.993
Term 2 not enrolled		1.100	1.063	1.084	1.049	1.036	1.111
Term 2 GPA (0.1s)		0.812 ***	0.852 **	0.871 *	0.878 *	0.884 *	0.876 *
Term 2 Hours Attempted		1.048 **	1.035 *	1.030	1.025	1.023	1.019
Term 2 Hours Earned		0.979	0.990	0.997	0.998	0.997	1.006
Term 3 not enrolled			1.163	1.006	1.025	0.995	1.007
Term 3 GPA (0.1s)			0.936	1.021	1.033	1.038	1.020
Term 3 Hours Attempted			1.069 ***	1.050 **	1.047 **	1.042 *	1.030
Term 3 Hours Earned			0.932 ***	0.949 **	0.953 **	0.955 *	0.974
Term 4 not enrolled				1.128	0.846	0.867	0.954
Term 4 GPA (0.1s)				0.846 **	0.895	0.910	0.908
Term 4 Hours Attempted				1.069 ***	1.052 **	1.051 *	1.049 *
Term 4 Hours Earned				0.925 ***	0.934 ***	0.935 ***	0.953 *
Term 5 not enrolled					1.657 *	1.490	1.471
Term 5 GPA (0.1s)					0.926	0.985	1.033
Term 5 Hours Attempted					1.062 **	1.047 *	1.035
Term 5 Hours Earned					0.958 *	0.971	0.995
Term 6 not enrolled						1.052	0.922
Term 6 GPA (0.1s)						0.868 *	0.922
Term 6 Hours Attempted						1.047 *	1.021
Term 6 Hours Earned						0.960	0.996

Table 6 (cont'd)

Post-Attendance

	Model 1		Model 2		Model 3		Model 4		Model 5		Model 6		Model 7		Model 8	
Not graduated from Commuter State															4.783	***
Total Loan Amount (1000s)															1.021	***
Intercept	0.112	***	0.097	***	0.080	***	0.067	***	0.062	***	0.051	***	0.055	***	0.014	***
N	13108		13108		13108		13108		13108		13108		13108		13108	

Overall Model Outputs	X^2	-2 Log likeli-hood	X^2	-2 Log likeli-hood	X^2	-2 Log likeli-hood	X^2	-2 Log likeli-hood	X^2	-2 Log likeli-hood	X^2	-2 Log likeli-hood	X^2	-2 Log likeli-hood	X^2	-2 Log likeli-hood
Omnibus Test of Model Coefficients	234	5525	327	5432	377	5382	420	5339	501	5258	525	5234	547	5212	717	5042
Hosmer & Lemeshow Test	4.814		7.466		14.249		13.520		20.570		10.727		12.777		11.153	
Cox & Snell R2	0.018		0.025		0.028		0.032		0.038		0.039		0.041		0.053	
Nagelkerke R2	0.050		0.069		0.080		0.089		0.106		0.110		0.115		0.150	

Note: Reference group or scaling within parentheses, significance level p<.001"***"; p<.01"**"; p<.05"*"

Pre-College Characteristics

Overall, the statistically significant pre-college characteristics in regression model 2.1 that did not change with the addition of the term enrollment variables (moving to models 2.2 and on) are sex, age, EFC, and first-generation status. The odds of defaulting are 1.3 times greater for males (p<.001). With each one-year increase in age when a student entered Commuter State, the odds of borrowers' likelihood to default increases by 1.6% (p<.05). Borrowers' who have $1000 more in EFC are approximately 3% (p<.001) less likely to default. First-generation status matters. The odds of defaulting are approximately 1.2 (p<.01) times greater for first-generation students when compared to non-first-generation peers. Further, the flag for missing admitted GPA became a significant predictor, beginning with model 2.3, with the addition of college experience by semester variables. In the next section, I discuss the changes to the college experience variables by introducing the term enrollment variables.

College Experiences by Semester Enrolled

The addition of the term enrollment variables increased the predictive ability of the models, although the increase in explained variation from regression model 2.1 to regression 2.2 was minimal. The addition of the variables did affect the college experience models changing, in some cases, which variables within each model are significant predictors of default. In the following sections, I discuss each college experience variable from Table 6.

STEM Major. The regression models in Table 6 show students' major matters. Specifically, for every college experience model, STEM major is statistically significant to, at least, a 0.05 p-level. Students studying a STEM major are less likely to default than their non-STEM peers. Depending on the model, the odds of defaulting for STEM majors are approximately 0.8 those of non-STEM majors.

110

Enrollment by term. Only one enrollment by term variable is a statistically significant predictor of default. Enrolling in term 5 (model 2.6) is the only enrollment variable significantly associated with default. Students who did not enroll in their fifth term are 1.7 (p<.05) times more likely to default than their enrolled peers. The remainder of the enrollment by term variables are not significant at the 0.05 level.

Term GPA. Term GPA still matters even when introducing the term enrollment variable. Changes in the early term variables remain significant for every college experience model with the enrollment flag (models 2.3 thru 2.7). The inverse relationship with defaulting stays the same, and the odds ratios are relatively similar.

Term hours attempted. The term hours attempted variable is statistically significant for every model that includes an enrollment flag (models 2.3 and on). Although the term enrollment variable is not a significant predictor of defaulting for each term, it helps amplify the importance of hours attempted by term. There is a positive relationship between hours attempted and defaulting for the models with the enrollment flag.

Term hours earned. Overall, when the term hours earned is a significant predictor, it shifted within the college experience models 2.3 thru 2.7 rather than an overall increase in the significance of the hours earned variables. Like term hours attempted, the direction of the relationship between term hours earned and defaulting remained the same. As such, with each additional hour earned per term, the odds of defaulting decreased.

In the chapter thus far, I discussed my findings related to my research questions. In the next section, referencing Table 7, I discuss my findings for each of the analytic analyses in my study (my three research questions) within the context of what the national default literature suggests are important findings. The "Xs" in the table indicate significant predictors of default

from national findings and for each of my three research questions. The purpose of this discussion and Table 7 is not as a comparison of my results to the national studies, but rather to highlight the context of the existing studies. The existing national studies focused on different contexts than my study and, therefore, there is a need to identify institution-specific correlates of default and not solely be guided by existing literature that focuses on different contexts. My results for research question one show which default factors are important for Commuter State students. The results show defaulters, on average, have more need, are more likely to be Pell-eligible, Black, have an unknown first-generation status, and be independent. This represents five variables specific to Commuter State that national studies find important predictors of default (Table 7).

Table 7. *The National Default Literature Findings and the Analytic Analyses Findings for this Study*

	Research Question 1		Research Question 2	Research Question 3
	National Findings	Descriptive Analysis	Regression Model 1: Traditional Default Model	Regression Model 2: Expanded Model with Term Enrollment
Pre-College Student Characteristics				
Male			X	X
Race/Ethnicity (White)	X	X		
Age	X			
Scaled EFC (1000s)	X	X	X	X
Pell-Eligible	X	X		
Scaled Commuter State Grant (1000s)	X		X	X
First-Generation Status (Non-First-Generation)	X	X	X	X
Independent Student	X	X		
Student has Dependents	X			
Rounded Mean Admitted GPA (0.1s)	X			
Missing admitted GPA			X	X
Mean ACT Composite	X			
Missing ACT Composite Score				
Post-Attendance				
Not graduated from Commuter State	X	X	X	X
Total Loan Amount (1000s)	X	X	X	X

For college experiences by semester enrolled, my results show differences in term GPAs and hours earned per semester between defaulters and those not defaulting. In addition, defaulters are more likely to be non-STEM majors. Finally, for the post-attendance variables, defaulters borrowed less and are less likely to have received a degree from Commuter State, both of which are important factors in national default studies.

My results for research question two show that the characteristics of Commuter State defaulters are generally mixed in relation to the national findings. Some characteristics of the findings align; however, other characteristics do not align. For regression model one, my research finds that four of the ten pre-college student characteristics identified in the previous default research findings are predictors of default at Commuter State (see Table 7). One of the most studied default variables, race/ethnicity, is not a statistically significant predictor of Commuter State defaulters. Beyond race/ethnicity, the academic preparation variables (incoming GPA and ACT composite score) are also not statistically significant predictors.

My regression model two for research question three shows institution-specific, student-level data at Commuter State increased the ability to predict default sooner. Regression model two provides a template to observe whether leaving at any point in time would predict defaulting and allow Commuter State administration to target stopped-out students as well as those who remained enrolled. As indicated in Table 7, four of the ten national study pre-college student characteristic variables are statistically significant predictors of default for my study of Commuter State students. The variables include sex, EFC, Commuter State grant aid, and first-generation status. The notable missing predictive variables from the model, which are significant in national studies of default, are race/ethnicity and academic preparedness (including incoming GPA and ACT composite score). The two post-attendance variables (graduation and total loan

114

amount) are statistically significant in regression model two and national studies. In the next chapter, I discuss the applicability of the research findings to Commuter State, how the findings could be a model for other institutions, and how national studies may be missing an essential component of understanding student default.

CHAPTER 5: DISCUSSION

My research examined the predictors of loan payment default that existing research has examined for other populations, including variables such as student characteristics (e.g., race, age, gender, socio-economic status, first-generation status), graduation status, the field of study, and total federal loan amount. As a result, my research expanded the existing understanding of loan default by incorporating institutional-level variables not traditionally captured in national datasets. My study expands what we know about default and provides practical implications by incorporating college experience by semester enrolled variables, such as term-specific GPA and hours attempted and earned. However, because my study was a quantitative case study for an exclusively commuter institution, my findings are not generalizable across all commuter students at all institutions. The advantage of my study's institution-specific approach is that it allowed me to include institution-level data into the models to help understand their contributions to defaulting. My approach allowed me to examine how institutional-level data can inform institutional decision making by showing which pre-college student characteristics, college experience by semester enrolled variables, and post-attendance factors associated with default for my subject institution. Further, my approach provided a process for other nonresidential institutions to examine the issue by providing a model to replicate to investigate default on their campus. Specifically, by incorporating these institutional-level variables in an organized way (i.e., by type of activity or by year), I can determine how early in a student's career I could identify them as "at-risk" of default. I begin this chapter with a summary of my findings. I then discuss the applicability of my research findings to Commuter State, how my findings could be a model for other institutions, and how national studies may be missing an essential component

about student default. I conclude with a discussion about my study's limitations and the opportunities for future studies to expand my research.

Research Findings Summary

The purpose of my study was to understand if institutional data could improve the timing of the estimates of default probabilities at a nonresidential campus. I summarize, in this section, my findings, discuss what they suggest, and unpack why they might be the way they are. My research questions organize my summary:

1) What are the characteristics of Commuter State students who do and do not default?

2) What predicts default among students at Commuter State?

3) Do institution-specific, student-level measures improve estimates of default for students at Commuter State?

The Descriptive Characteristics of Commuter Students at a Nonresidential Campus who Default

The purpose of research question one was to identify the descriptive characteristics of commuter students at a nonresidential campus who do and do not default on their loans. Commuter State students who defaulted, on average, had more need and borrowed less. Further, the students who defaulted were more likely to be Pell-eligible, Black, independent, an unknown first-generation status, and not graduated from Commuter State. The odds of defaulting for Commuter State students with an unknown first-generation status was almost twice as likely compared to their non-first-generation peers. My findings for the first-generation variable may be a result of the missing first-generation data at Commuter State. My results for the study could be different if Commuter State had complete first-generation data for their students, therefore eliminating or, at least, minimizing the number of unknown first-generation status students.

117

There were no notable differences in academic preparation measures between the not defaulted and defaulted students who attended Commuter State. The general trend is that the higher the incoming admission GPA and standardized test score, the lower the likelihood of default. My findings for the academic preparation variables suggest that default is not associated with prior academic preparation. This may be a result of the small overall variation of incoming preparation of the students. In other words, Commuter State's admission standards may limit the variation of the student's incoming academic preparation and, therefore, by limiting the academic preparation variation could affect the measurable association between academic preparation and default. Another potential reason differences did not emerge with the academic preparation measures is the large volume of missing incoming GPA and composite test scores in the data set. The predictive ability of the academic preparation measures could increase if the data set had fewer missing fields in the academic preparation measure. In this section, I outlined for research question one my findings for the descriptive analyses of the key characteristics for who do and do not default on their loans from Commuter State. . . In the next section, I discuss my findings for research question two.

Predicting Default at Commuter State Using the Traditional Measures

My results for research question two model the characteristics that predict default for Commuter State students. One of the most studied default variables, race/ethnicity, was not a statistically significant predictor of Commuter State defaulters. Interestingly, my research question one findings show that Black students are more likely to default than their White peers. However, due to the nature of descriptive analyses methodology, my results for research question one only look at differences for each variable independent of the other variables. For research question two, because I utilized regression methodology, my findings show the

118

significant predictors of defaulting when controlling for the other variables in the model. When I loaded all the variables into the regression model, race/ethnicity's association with default diminished. The variable was not a significant predictor of default for Commuter State students. My findings suggest that when accounting for all the variables in regression model one, there are other pre-college student characteristics with a greater association with default than the students' race/ethnicity. My results indicate that Commuter State students' sex, incoming SES, the amount of Commuter State grant aid received, and first-generation status are the important characteristics associated with default.

Unlike the race/ethnicity variable, the academic preparation variables were not statistically significant predictors of default for research question two, which aligned with the findings from research question one. My findings suggest that the other significant predictors of default are the important factors for Commuter State students. From Table 7, missing admitted GPA was a statistically significant predictor of default. If my data set had fewer missing admitted GPAs, would I get different results for the admitted GPA variable?

Institution-Specific, Student-Level Measures Improve Estimates of Default for Students at Commuter State

My findings for research question three demonstrate that institution-specific, student-level data at Commuter State provides a template to predict default sooner beyond the traditional default measures. My results suggest the existing default research may be missing an important set of variables in their models, variables measuring students' college experiences. Although my findings are unique to Commuter State, future national studies should include students' college experiences in the models to measure whether college experience variables increase the

119

understanding of default at a national scale. The greater our ability to understand the predictors of default, the greater our ability to mitigate students' likelihood to enter into default.

Specifically for Commuter State, my findings provide specific pre-college student characteristics, college experience factors, and post-attendance variables for Commuter State administrators to consider when determining which students are likely to default. My findings can give a semester-by-semester ability to predict default or a model to predict default for students who have exited Commuter State and begun their loan repayments. In the following section, I discuss the applicability of my research findings for Commuter State.

The Applicability of the Research Findings to Commuter State

The results from my study provide multiple findings for Commuter State. Specifically, my results could inform admissions policy, financial aid policy, and student success. In the following sections, I discuss the applicability of my findings to Commuter State for admissions, financial aid, and student success.

Applicable Research Findings for Admissions Policies

My results show the key characteristics associated with defaulting for Commuter State's population. These findings can help inform data driven admissions policy development at Commuter State in two ways. First, the findings inform the need to consider changing the process of collecting students' first-generation status. Second, my results may also help inform the transition to more holistic metrics for admitting students.

The traditional characteristics of the defaulter in the national models were not all predictors of defaulting for the Commuter State students. Table 7 shows that this holds for all three of my research question results. Specifically, the first-generation status variable is a traditional pre-college student characteristic associated with default and was a predictor of

defaulting for Commuter State students. However, by looking at the descriptive analysis results, I found a unique difference with applicable implications for the Commuter State administration. Via the descriptive analysis results, Commuter State defaulters, on average, were more likely to have an unknown first-generation status. My results found that the odds of the unknown group defaulting were two times greater than non-first-generations students. As outlined in chapter 4, the Commuter State unknown group are the students who did not provide first-generation data. These findings suggest it is important to implement policies and practices to ensure that Commuter State collects first-generation data for every incoming student. Such a practice could be a required first-generation question on the admissions application. Every student has to complete the admissions application before they are considered for admission, and, therefore, a required question on the application would ensure Commuter State has the first-generation status for every enrolled student. My three regression models found that first-generation is a statistically significant predictor of default. However, would the results change if Commuter State enacted the above policy that collects first-generation status for every enrolled student? Would my findings change if approximately 5% of the students with a missing first-generation status moved to have a status? For instance, would the first-generation variable predict default if most of the missing status students moved to a first-generation status? What if a large proportion of the missing status students moved to non-first-generation status? Would the first-generation variable be insignificant? Having fewer students without a missing status would provide a more detailed snapshot of the association of first-generation status and default.

Beyond first-generation status, my findings could inform admissions policies related to Commuter State's criteria for admissions. The results of the descriptive analysis and regression models did not yield distinct differences for Commuter State defaulters from their not defaulting

peers with regards to typical incoming admissions measures, including GPA and standardized test scores (measured by ACT composite score for my study). These findings provide additional information to the Commuter State administration when considering admission policy changes. For example, many institutions adjusted their admissions policy to test-optional for the entering classes during the worldwide pandemic in 2020 (Rhyneer, 2021). Over 700 institutions adopted a test-optional policy in 2020; this was five times more institutions in one year than in the previous four years combined (Rhyneer, 2021). Commuter State removed the required submission of a standardized test score to complete an admissions application. If Commuter State decides to review the viability of a permanent test-optional admissions policy, the leadership would consider many factors related to the effectiveness of requiring a standardized test score to be admitted. My results show composite ACT scores were not associated with default for two of the three regression models and the descriptive analysis results. Commuter State's administration could use my findings as additional information in reviewing the viability of a test-optional policy at Commuter State.

Applicable Research Findings for Financial Aid Policies

Beyond informing admissions policy at Commuter State, my research findings could also provide context to Commuter State's financial aid strategy. The findings for the variables EFC and Commuter State Grant were significant predictors of default for all three of the regression models. My findings for EFC show that the wealthier a background a student comes from, the less likely they are to default. While the results from my regression models also found that as the amount of grant aid allocated to students from Commuter State increased, the likelihood of default decreased. These results confirm the importance of aid to reduce out-of-pocket tuition and fees costs for students. Commuter State administration could consider allocating more

122

financial aid to lower EFC students, hopefully helping mitigate the students' debt burdens further. Another way my findings could inform financial aid policies is by restructuring the terms and conditions associated with receiving grant aid from Commuter State. My results show that lower EFC students, those who are needier, are more likely to default. Utilizing my findings regarding needier students and default, Commuter State could target low EFC students with additional need aid. The terms and conditions of receiving this aid could require that these students complete financial literacy training provided by Commuter State each year. Part of the training could include teaching the students about the available Federal Student Loan repayment plans. The policy provides intervention for at-risk borrowers to reduce their barriers to enrolling in affordable repayment plans. This type of policy is an example that aligns with research from The Pew Charitable Trusts (2020) that points to actions the Department of Education and Congress can do to help borrowers avoid default, including identifying at-risk borrowers and eliminating barriers to enrollment in affordable repayment plans. This is important because institutions have a vested interest in keeping their 3-year default rates low. Individual institutions risk losing eligibility for federal student aid programs if their 1-year and multi-year default rates exceed pre-defined thresholds set by the Department of Education (NASFAA, 2021). Since my study provides a method to allow campuses to use local data to identify potential defaulters and take action, it provides an avenue to minimize default and, thus, can help reduce individual institutions' overall default rates.

Applicable Research Findings for Student Success

Commuter State is committed to student success. One of the four pillars of its university strategic plan is student success. As such, my findings provide applicable actions for Commuter State to apply to their strategies to see students succeed. My results show that graduation is the

strongest predictor of Commuter State students not defaulting. Commuter State can apply this finding as another reason to embed getting students to graduation in every student success practice. For example, when advisors advise students, they can share that students who graduate are more likely to repay their loans than those that do not. Another example is that Commuter State could incorporate the increased likelihood of defaulting for students who do not graduate within their publications.

Beyond graduation, my findings provide a practical semester-by-semester prediction model to inform university administration of students at risk of default. The administration could implement the findings from my model to calculate a default likelihood for every enrolled student each semester. University advising could use the results to create individualized enrollment plans to ensure students minimize their likelihood of default by enrolling in a prescribed amount of credits. Further, university support offices could utilize the individually calculated default likelihood to identify which students, measured by their term GPA, should engage with university tutoring and academic support services to mitigate future default.

My study's findings have applicable applications to inform numerous policies and interventions practices at Commuter State. Moreover, my findings apply to numerous areas of Commuter State supporting admissions policy, financial aid strategy, and student success practices. The wide array of practical implications of my study's findings for Commuter State shows the importance of studying student loan default as a case study at one institution.

Practical Default Model for Other Institutions

The previous section outlined how my findings have practical implications for admissions, financial aid, and student success policies and practices at Commuter State. In this section, I build upon my discussion from the previous section to share how other institutions

could use my findings for their institution-specific context. Any institution could replicate my study with its institutional data. Almost all, if not every, variable I used in my regression models are data points institutions have and can access for their students. The only real caveat is that the institution must receive federal financial aid funding and, therefore, their students would receive federal student loans. As long as the institution receives federal financial aid, they would have all the variables I include in my models that are variables received from the FAFSA. Beyond the FAFSA data, the pre-college, student experience, and post-attendance data I utilized in my models are traditional metrics many institutions capture through their admissions, enrollment, and financial aid reporting processes.

I have established that other institutions would have the necessary data to replicate my model at their campuses. The next step is to discuss how to structure the models. Within my study, I specifically outline which variables load into which blocks and discuss the order to load the blocks of variables into the model, which provides a roadmap for other institutions to replicate the loading and analyzing I have done in my study with their institution-specific data. Institutions could utilize my model to generate their institution-specific findings to inform practical changes to policies and procedures at their campuses to mitigate their students from defaulting on their federal student loans.

The Missing Component from National Studies

National studies may be missing an essential component to our understanding of student default. The traditional default model includes pre-college student characteristics and post-attendance variables. My study expanded the traditional model by adding college experience factors and my research results found that adding these blocks of variables increased the existing understanding of default. Specifically, my study shows that local, institutional-specific data

125

increases the ability to predict default earlier by increasing the model's explained variation. My study is a template that offers a way to get default estimates sooner before a borrower defaults.

Regression model two, which included the college experience blocks of data in addition to the traditional pre-college and post-attendance blocks of variables, show that local data increased the ability to predict default for students at Commuter State earlier than existing research. These findings beg the question, if national studies' included college experience variables, how would the national model results change? Could national models predict default sooner? Would the traditional measures of default established in national studies change by adding the college experience blocks? My results from regression model3, shown in Table 7, did find that some of the traditional predictors of default from national studies were not predictive in my complete model for regression model two, which included the pre-college student characteristics, the college experience blocks, and the post-attendance variables. My findings support the need for institution-specific default research if institutions want to increase their understanding of default for their own students.

In the chapter thus far, I outlined how my findings have applicable applications for Commuter State's policies and practices related to admissions, financial aid, and student success. I discussed how my study provides a practical default model for other institutions to replicate, and I outlined how my findings may provide a missing component to national studies. In the following section, I transition from discussing my findings to discussing some of the limitations of my study.

Limitations

My study has limitations. It is limited by only studying one institution, the 3-year default window, and the lack of available post-attendance data. My findings are limited because they are

not generalizable across all nonresidential institutions due to only studying students at one institution. Further, my study is limited to the 3-year default data provided by the Department of Education. Finally, the lack of available post-attendance data (i.e., employment status, earnings, life situations, other debt such as cars, homes, and credit cards) for Commuter State students limited my study. In the following section, I discuss these limitations in further detail.

The findings from my research are not generalizable. Focusing on one institution, Commuter State, I established that Commuter State is different from national default findings. However, I cannot state that my findings are generalizable for all commuter students by only studying Commuter State. Although my study does not provide generalizable default findings, it does provide a model that other institutions can replicate to understand defaulting on their campuses.

The available default data limit my study. The Department of Education default data only track the first three years of borrowers' repayment and, therefore, I do not measure the complete debt situation. Borrowers are likely to default beyond their first three years of repayment. In my study, I do not measure any default beyond the 3-year window; therefore, my default measurement may be relatively low. Would my findings change if the default measurement was a longer window, such as five or ten years? If the default measurement extended, would my models' total explained variation decrease substantially because college enrollment was further away and more post-college situations could occur that would positively or negatively affect default?

Beyond the limitation of the default measurement, the lack of available post-attendance variables limits my study. It is limited by not capturing borrowers' complete debt obligation. For example, my models did not include private and personal loans such as car, credit card, and

potentially mortgage debt. The study should include all of the borrowers' debt obligations to understand the complete picture of their ability to repay their federal student loans successfully. My study does not include employment status nor earning data because this data is not available at Commuter State's SIS nor from the NSLDS data sets. These are both important factors related to borrowers' ability to repay their loans. Moreover, I cannot discern whether the borrowers' have engaged in a repayment plan, such as an income-contingent repayment plan that caps the total monthly payment based on borrowers' monthly income. This is important because research from The Pew Charitable Trusts (2020) suggests income-contingent repayment plans help mitigate the risk of borrowers defaulting on their repayment.

Ways to Expand My Study

My study uniquely contributes to the existing default research. It provides an understanding of the key characteristics associated with students' default at Commuter State. My study provides a model for other institutions to replicate to understand the nuances of defaulting of their students. However, as noted in the previous section, my study has limitations. These limitations are also opportunities to expand my research. One could expand my study by adding multiple nonresidential campuses into the data set. The findings from a multi-institution study would provide a greater understanding of whether commuter student default is unique to Commuter State or representative of the larger commuter population. One could further expand my study by accessing more granular default data. This would require working with the Department of Education to explore if they systematically record defaults beyond the 3-year window. If so, collecting this expanded default window measurement (i.e., 5- or 10-year default window) would expand my study by bringing a complete picture of who defaults over a longer timeframe than 3-years after exiting higher education. Finally, expanding the post-attendance

128

block of variables to include additional debt obligations, employment status, and net wealth, including borrowers' earnings and investments, would expand the measurable footprint of the factors potentially contributing to default after students exit higher education.

Conclusion

My study aimed to understand the defaulted student characteristics from a nonresidential college. I examined if adding institution-specific, local data increased the ability to predict default sooner. My results show that Commuter State's predictors of default for nonresidential students differ from national default studies, supporting the need to conduct institution-specific default research. My study shows that local data, specifically college experience by semester enrolled data, contributes to the overall ability to predict default semester-by-semester instead of waiting until students graduate or leave the institution. My findings provide an applicable default template Commuter State administration can use on their campus for data driven decision-making. Further, my study provides a model other institutions could replicate for their campuses. With over a trillion dollars in outstanding Federal Student Loan debt, loan default is an important topic, and my study contributes to the existing default research.

APPENDICES

APPENDIX A: Data Preparation Process

Data preparation was a multi-step process**.** The data was cleaned and transformed into a structure (format) that met the modeling requirements. I conducted a data analysis process to review the data to identify and address errors and abnormalities that existed within the data set. The analysis identified missing data as well as abnormal results. In order to determine if the missing data skewed the analysis, robustness checks with and without imputed data were conducted.

Concerning data abnormalities, errors can and do happen when data is stored in the subject university's student information system (SIS). For example, some data fields are manually entered, which can result in accidental augmented syntax, leading to errors. For instance, a student could have a listed GPA of 304 instead of 3.04 with an accidental omission of the decimal. Simple descriptive techniques were utilized to check for such errors. Further, aggregate results from the descriptive statistical analyses were validated against aggregate institutional reporting. Any large discrepancies were investigated further to ensure the difference was valid or uncover if there was a deeper issue with the data that was extracted from the SIS.

I conducted regression model B1 as a robustness check. I begin this section by discussing

the similarity of regression models B1 and two. I then follow with a discussion of how the

models contrast and the reason for the two approaches for the same research question. I conclude

with a discussion of my findings for regression model B1.

Similarities of Regression Models B1 and Two

Regression models B1 and two utilized all three blocks of data but in a slightly different

configuration. I outline the similarities of the model in this section and discuss the differences in

the following section. The *a priori* sequence of loadings the blocks of variables into regression

model B1 and two were the same. The pre-college student characteristics block was the first

block of variables added to both regression models. After fitting the association between pre-

college characteristics and default, the next set of blocks of predictors loaded into the models are

the factors measuring college experiences by semester enrolled. The final block of variables

loaded into regression models B1 and two are the post-attendance variables. The decision to load

the post-attendance variables was to account for one of the most important predictors of

defaulting, whether a student graduated or not. Loading graduation into the model as the last

block explained how influential graduation and debt were to defaulting after accounting for the

other variables in the model.

The Difference of Regression Models B1 and Two

I outline in this section the difference between regression models B1 and two. It is typical

to see students drop out semester over semester. For those students who dropped out, their term

data is blank. For instance, every student in the data set was enrolled in term 1. Beginning with

term 2, students started dropping out; thus, these students' data for term 2 is missing, so the

regression model's N decreases to only those students still enrolled, those students who have term 2 enrollment data. Therefore the model only provides results for the students still enrolled, not taking into account the students who have dropped out. As I move from term 2 to 3 and so on, the number of students still enrolled, and thus included in the model, continues to decrease. Regression model B1 tells me all the students who are still enrolled semester by semester and will continue to enroll semester by semester. It does not capture the students who dropped out. This is why I developed regression model two. My regression model two expands regression model B1 by including both the enrolled and dropped out students.

Ultimately, regression model B1 only tells me about the predictors of default for enrolled students each term, those who persisted. The value of this model is its ability to provide specific enrollment characteristics by term that are related to default for those students still enrolled. In other words, regression model B1 provides a default intervention model specific to Commuter State for those students who are still enrolled but provides no information on drop-outs until the final completion block is added potentially many years down the line—which ultimately may only contain information on those students who dropped out after six semesters of attendance (since all the other students have been omitted term by term). Regression model two expands regression model B1 by providing default predictions for both the enrolled and not enrolled students at each point in time, measured by terms.

Regression Model B1 Results

In this section, I discuss regression model B1, which utilizes the variables included in the data from Commuter State. The unique contribution regression model B1 provides to my study is an understanding of predicting future default for currently enrolled students. The value of this model is its ability to provide specific enrollment characteristics by term that are related to

133

default for those students still enrolled. In other words, regression model B1 provides a default

intervention model specific to Commuter State. Overall, after controlling for pre-college

characteristics, the addition of the college experience category helped explain additional

variation within the model. Specifically, the logistic regression analyses show that the explained

variation increases incrementally with each additional block of college experience variables

entered into the model. My model results confirm institution-specific, student-level data for

Commuter State students matter, contributing to our ability to predict default. In the following

section, I begin by discussing the pre-college characteristic findings for the regression model. I

then discuss how the explanatory ability of the model changes as blocks of term enrollment

variables are entered into the model.

Pre-College Characteristics

Model B1.1: pre-college characteristics shows the logistic regression results for

defaulting at Commuter State (see Table 8). My results in model B1.1 show sex, race/ethnicity,

age, EFC, Pell-eligibility, first-generation status, and admitted GPA are statistically significant

predictors (p<.05) of defaulting. The odds of defaulting are 1.3 times greater for males (p<.001),

1.5 times greater for Black borrowers (p<.001) when compared with Whites. In addition, with

each one-year increase in age when a student entered Commuter State, the odds of borrowers'

likelihood to default increases by 1.6% (p<.05). Borrowers' who have $1000 more in EFC are

3.1% (p<.001) less likely to default. The odds of defaulting are 1.3 (p<.05) times greater for Pell-

eligible students when compared to their non-Pell-eligible peers. First-generation status matters.

Model B1.1 shows the odds of defaulting are 1.2 (p<.01) times greater for first-generation

students when compared to non-first-generation peers. Lastly, each decrease of 1 point in

admitted GPA increased the odds of defaulting by 32% (p<.001). The aforementioned pre-

college student characteristics are the statistically significant factors that predict defaulting when only considering pre-college characteristics in the model. I discuss in the next section the STEM major variable then transition to a discussion of how adding college experience by term factors into the regression changes the ability to predict default at Commuter State.

Table 8. *Regression Model B1: Logistic Regression Estimates Post-Attendance, Pre-College Student Characteristics, and College Experience by Semester Enrolled (Expressed in Odds-Ratios)*

	Model B1.1: Pre-College Student Characteristics		Model B1.2: College Experience Semester 1		Model B1.3: College Experience Semester 2		Model B1.4: College Experience Semester 3		Model B1.5: College Experience Semester 4		Model B1.6: College Experience Semester 5		Model B1.7: College Experience Semester 6		Model B1.8: Post-Attendance	
Pre-College Student Characteristics																
Male	1.301	***	1.289	**	1.238	*	1.236	*	1.238	*	1.293	*	1.210		1.154	
Race/Ethnicity (White)																
Black	1.497	***	1.264	*	1.220		1.227		1.118		1.072		1.143		0.963	
Hispanic	1.329		1.229		1.132		0.996		0.908		0.938		1.023		0.911	
Other URM	1.038		1.016		1.037		0.998		1.009		1.017		0.990		1.045	
Unknown	1.286		1.231		1.002		0.882		0.967		0.978		1.401		1.493	
Age	1.016	*	1.017	*	1.021	**	1.024	**	1.031	***	1.037	***	1.028	**	1.013	
Scaled EFC (1000s)	0.969	***	0.973	***	0.973	***	0.974	**	0.971	**	0.973	**	0.976	*	0.978	*
Pell-Eligible	1.304	*	1.189		1.169		1.092		0.965		0.947		0.859		0.839	
Scaled Commuter State Grant (1000s)	1.005		1.028		1.025		1.020		1.022		1.028		1.042		1.052	*
First-Generation Status (Non-First-Generation)																
First-Generation	1.239	**	1.222	*	1.170		1.265	*	1.287	*	1.277	*	1.395	**	1.340	*
Unknown	1.266		1.246		1.296		1.235		1.186		1.102		1.225		1.227	
Independent Student	1.156		1.150		1.112		1.068		0.977		0.898		0.964		0.868	
Student has Dependents	0.683		0.681		0.665		0.565	*	0.520	*	0.558		0.635		0.630	
Rounded Mean Admitted GPA (0.1s)	0.684	***	0.823		0.918		0.849		0.902		0.808		0.729		0.758	
Missing admitted GPA	1.115		1.157		1.169		1.272	*	1.386	**	1.515	***	1.563	***	1.750	***
Mean ACT Composite	0.990		0.997		0.990		0.982		0.978		0.962		0.949	*	0.923	**
Missing ACT Composite Score	1.035		1.088		1.158		1.234		1.331	*	1.330	*	1.389	*	1.547	**

Table 8 (cont'd)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
College Experiences by Semester Enrolled								
STEM Major		0.802 *	0.759 **	0.750 **	0.720 **	0.700 **	0.691 **	0.718 *
Term 1 GPA (0.1s)		0.783 ***	0.858 **	0.854 **	0.874 *	0.921	0.965	0.980
Term 1 Hour Attempted		1.029	1.000	0.990	0.988	1.007	0.998	0.989
Term 1 Hours Earned		0.968 *	0.982	0.999	1.005	1.001	1.003	1.018
Term 2 GPA (0.1s)			0.807 ***	0.867 *	0.881	0.900	0.896	0.897
Term 2 Hours Attempted			1.056 **	1.034	1.037	1.016	1.022	1.011
Term 2 Hours Earned			0.978	0.997	1.003	1.009	1.011	1.029
Term 3 GPA (0.1s)				0.934	1.028	1.040	1.049	1.038
Term 3 Hours Attempted				1.078 ***	1.043 *	1.035	1.026	1.007
Term 3 Hours Earned				0.927 ***	0.951 *	0.962	0.967	0.991
Term 4 GPA (0.1s)					0.830 **	0.873 *	0.881	0.890
Term 4 Hours Attempted					1.084 ***	1.055 *	1.045	1.039
Term 4 Hours Earned					0.920 ***	0.933 **	0.939 **	0.953 *
Term 5 GPA (0.1s)						0.930	1.021	1.061
Term 5 Hours Attempted						1.074 ***	1.069 **	1.056 *
Term 5 Hours Earned						0.952 *	0.952 *	0.974
Term 6 GPA (0.1s)							0.863 *	0.919
Term 6 Hours Attempted							1.061 *	1.035
Term 6 Hours Earned							0.956	0.993
Post-Attendance								
Not graduated from Commuter State								5.393 ***
Total Loan Amount (1000s)								1.020 ***
Intercept	0.112 ***	0.097 ***	0.081 ***	0.102 ***	0.075 ***	0.090 **	0.160 *	0.066 **
N	13108	13108	12350	11209	10482	9805	9098	9098

Table 8 (cont'd)

Overall Model Outputs	X^2	-2 Log likeli-hood	X^2	-2 Log likeli-hood	X^2	-2 Log likeli-hood	X^2	-2 Log likeli-hood	X^2	-2 Log likeli-hood	X^2	-2 Log likeli-hood	X^2	-2 Log likeli-hood	X^2	-2 Log likeli-hood
Omnibus Test of Model Coefficients	234	5525	327	5432	331	4894	295	4170	322	3606	267	3170	249	2808	419	2638
Hosmer & Lemeshow Test	4.814		7.466		11.456		15.335		15.006		8.082		12.292		3.015	
Cox & Snell R2	0.018		0.025		0.026		0.026		0.030		0.027		0.027		0.045	
Nagelkerke R2	0.050		0.069		0.077		0.079		0.097		0.091		0.094		0.158	

Note: Reference group or scaling within parentheses, significance level p<.001"***"; p<.01"**"; p<.05"*"

138

STEM Major

The regression models show students' major matters. Specifically, for every college experience model, STEM major is statistically significant to, at least, a 0.05 p-level.[11] Students studying a STEM major are less likely to default than their non-STEM peers. The odds of defaulting for STEM majors is 20% (p<.05) to 31% (p<.01) less likely than non-STEM majors, depending on the model.

College Experiences by Semester Enrolled

In this section, I discuss the regression results for adding blocks of college experience factors into the regression model in addition to the already loaded pre-college student characteristic variables. Overall, my results show institution specific, student-level data increases the ability to predict default. The explained variation increases with each additional set of term-specific hours attempted, hours earned, and term GPA loaded into the model. Six semesters of data were entered into the model, represented in Table 8 in Models B1.2 thru B1.7.

The Nagelkerke R^2 values increased incrementally, with each additional block of term college experience variables entered into the regression analysis, with the only exception being the addition of the term 5 block of variables. The addition of the term 5 variables decreased the explained variation by 0.006 (Nagelkerke $R^2 = 0.091$). However, the addition of term 6 variables increased the model's explained variation again. The general increase in the Nagelkerke R^2 value with each additional block of term college experience variables indicates that the understanding of defaulting increases as additional local, institutional-specific data is entered into the model. My results imply that as I add more semesters of college experiences into the model, the more complete of a snapshot of the students' enrollment the college experience data is capturing, and,

---

[11] STEM major is the classification of the students' major the last term they were enrolled.

thus, the greater the ability to predict default. This is an important finding because it shows how local data can improve the ability to predict default and expand previous default models.

I discussed in the previous section how the college experience factors and the overall model results changed as additional blocks of college experience variables were entered into the regression model. In the following section, I discuss the results of the college experience variables in greater detail. After discussing the college experience model results, I analyze how the pre-college student characteristics set of variables changed with the addition of the college experience variables. I did this to see if there are correlations between the college experience variables and the pre-college variables or if adding these variables explains additional variation. Suppose the pre-college estimates are stable with the addition of the new variables and the overall model fit increases. In that case, I can be confident that the additional variables explain the new variation and not simply pull variation previously explained by the pre-college variables.

My results answer research question three, finding institution-specific, student-level data increases the ability to predict default. My findings are confirmed via the increases in explained model variation with each additional block of term variables entered into the model.

Term GPA. The first college experience variable, term GPA, is a significant predictor of defaulting overall. For four of the six college experience models, the most recent term GPA is a significant predictor of default, while the previous term GPAs in the models would lose their predictive significance, or the odds ratio would get closer to 1. Moreover, in every model with a significant term GPA variable, term GPA has an inverse relationship to the odds of default. As term GPA increased, the odds of defaulting decreased.

Overall, it is not important to analyze how the term GPA's predictive significance changed with each additional block. Instead, it is important to note that the term GPA adds

additional information to the model and is in the correct direction. First, with each additional term GPA added to the regression model, some additional information is explained, measured by an increase in the explained variation from models B1.2 thru B1.7. Second, for each statistically significant term GPA in the models, the direction of the relationship between term GPA and defaulting is important. My results show that as term GPA increases, the probability of defaulting decreases for statistically significant terms.

Term hours attempted. The second college experience variable investigated in my research was credit hours attempted by term. The results show that the semester hours attempted variables that were later into the students' enrollment at Commuter State mattered, and when the variables are statistically significant in the models, they have a positive relationship with the odds of default. With each additional hour attempted per term, the odds of default increased.

The early terms did not matter. The term 1 hours attempted variable is not statistically significant in any of the six college experience models. The term 2 hours attempted variable is only significant in one of five of the college experience models. The terms 4, 5, and 6 hours attempted variables are more often predictors of default; however, their odds ratios are relatively close to one, ranging from 1.055 to 1.084.

Term hours earned. The final college experience variable was term hours earned. Like hours attempted, the middle to late terms are the significant predictors of default. Overall, the trend is an important take-away from my models. My results show that the more credits earned in a given semester, the lower the odds of default. The more credits earned, the closer a student moves to graduation, which aligns with Podgursky et al.'s (2002) findings that progress toward a degree is a significant predictor of not defaulting.

In the previous section, I discussed my findings that college experience factors help explain defaulting beyond the variation explained by the pre-college student characteristics. The Nagelkerke R^2 values increased incrementally, with each additional block of term college experience variables entered into the regression analysis, with the only exception being the addition of the term 5 block of variables. The addition of the term 5 variables decreased the explained variation by 0.006 (Nagelkerke $R^2 = 0.091$). However, the addition of term 6 variables increased the model's explained variation again. In the previous section, I discussed how the college experience factors and the overall model results changed as additional data were loaded into the model. However, what happens to the pre-college student characteristics when college experience variables were entered into the model?

Changes to Pre-College Characteristics as a Result of Adding College Experience and Post-Attendance Blocks of Variables

In this section, I address how the significance of the pre-college student characteristics changed in models two thru seven as the additional sets of term data were added. Overall, pre-college student characteristics remained statistically significant with the inclusion of institution specific, student-level data from Commuter State into the models. However, the pre-college factors that mattered changed as additional terms were introduced. Moving to the entire model, model B1.8: Post-Attendance in Table 8, which includes all the college experience and post-attendance variables, my research finds shifting occurs in which pre-college student characteristics predict default. Race/ethnicity, Pell-eligibility, and admitted GPA became statistically insignificant. Conversely, the variables Commuter State grant aid, the flag for missing admitted GPA, ACT composite score, and the flag for missing ACT composite score became statistically significant predictors of defaulting. Within this section, I discuss how the

statistically significant pre-college factors in model B1.1 changed in models B1.2 thru B1.7. I then discuss how some of the non-statistically significant variables in model B1.1 changed in models B1.2 thru B1.7.

Of the statistically significant variables in model B1.1, age and EFC remain significant in every college experience model (models B1.2 thru B1.7). Sex remains a statistically significant predictor of defaulting until the last set of term variables were added in model B1.7. First-generation status is similar to the sex variable in that it is significant for all but one of the college experience models. However, unlike the variable sex, first-generation status's significance value went above the .05 threshold for only one model but became significant again. In model B1.2, first-generation status is statistically significant (p<.05), it rose above the significance threshold in model B1.3 and became statistically significant again at a .05 threshold or less for models B1.4 thru B1.7. The statistically significant variables of Pell-eligibility and admitted GPA in model B1.1 are not significant predictors of default in any college experience models. Somewhat similar to Pell-eligibility and admitted GPA, race/ethnicity remains statistically significant for the first model with college experience factors, model B1.2, and then is not statistically significant for the remainder of the college experience models. The results from my regression analyses find Pell-eligibility, admitted GPA, and race/ethnicity, statistically speaking, do not matter as they relate to defaulting once institution-specific, student-level enrollment variables are introduced into the model. I discussed how the statistically significant variables in model B1.1 change due to introducing college experience sets of factors into the model. In the following paragraph, I discuss the pre-college variables that change due to adding the college experience variables.

Four variables that are not significant predictors of default (p<.05) in the pre-college student characteristics model B1.1 became significant for at least one model when I added the college experience set of factors into the regression model. The variables student has dependents, dummy for missing admitted GPA, ACT composite score, and the dummy for missing ACT composite score are the four factors that became significant in at least one model. In models B1.4 and B1.5, students who had dependents are more likely to default than their peers. The odds of defaulting are 1.4 (p<.05) and 1.5 times (p<.05) greater for students with dependents compared to their Commuter State peers, respectively. Unlike the student has dependents variable, once the dummy for missing admitted GPA is a significant predictor of defaulting, it remained significant for the rest of the college experience models. Students who did not have an incoming GPA are more likely to default than their peers in models B1.4 thru B1.7. Not only are they more likely to default, but their odds of defaulting increased with each additional set of term variables added for models B1.4 thru B1.7. Incoming ACT composite score is not a significant predictor of defaulting in the pre-college student characteristics model B1.1; however, it is significant in model B1.7, the final college experience model. With each one-point increase in composite ACT score, the odds of borrowers' likelihood to default decreased by 5.1% (p<.05). The fourth pre-college student characteristic that is not significant in model B1.1 and became significant is the dummy for the missing ACT composite score. This variable is significant in the last three college experience models, models B1.5 thru B1.7. Similar to the dummy for missing admitted GPA, once the dummy variable for missing ACT composite score is statistically significant, it remained significant, and its odds ratio increased with each addition of blocks of term variables.

Overall, the results from Commuter State confirm, pre-college student characteristics matter to default. Going a step further and adding institution-specific, student-level variables

144

shows estimates of defaulting can increase. College experience variables increase our ability to predict default and negate some of the predictive significance of pre-college characteristics, showing that college experiences matter. Although the results in Table 8 confirm college experiences are associated with defaulting, the results do not show whether it is the specific term variables of GPA, hours attempted, and hours earned that are associated with defaulting or if something else is happening, such as enrolling or not enrolling in each of the specific terms. In regression model two, I discuss my regression findings of incorporating term enrollment into the models.

REFERENCES

146

REFERENCES

Akers, B., & Chingos, M. M. (2016). *Game of loans: The rhetoric and reality of student debt.* Princeton University Press.

Allison, P. D. (2002). *Missing data*. SAGE Publications, Inc. https://www.doi.org/10.4135/9781412985079

Baum, S., Ma, J., & Payea, K. (2012). *Trends in public higher education: Enrollment, prices, student aid, revenues and expenditures.* College Board, Advocacy & Policy Center.

Baum, S., Ma, J., Pender, M., & Bell, D. (2015). *Trends in student aid 2015.* The College Board.

Baum, S., Ma, J., Pender, J. M., & Welch, M. (2017). *Trends in student aid 2017.* The College Board.

Ma, J., & Pender, M. (2021). *Trends in college pricing and student aid 2021.* The College Board.

Biddix, J. (2015). Editor's Notes. In J. P. Biddix (Ed.), *New directions for student services*: no. 150, Summer 2015. *Understanding and addressing commuter student needs* (pp. 1–2). https://doi.org/10.1002/ss.20122

Blagg, K. (2018). *Underwater on student debt: Understanding consumer credit and student loan default*. Urban Institute. https://www.urban.org/sites/default/files/publication/98884/underwater_on_student_debt. pdf

Bound, J., Lovenheim, M., & Turner, S. (2010). Why have college completion rates declined? An analysis of changing student preparation and collegiate resources. *American Economic Journal: Applied Economics*, *2*(3), 129–157. https://doi.org/ 10.1257/app.2.3.129

Burdman, P. (2005). *The student debt dilemma: Debt aversion as a barrier to college access*. Center for Studies in Higher Education, the University of California.

Carnegie Foundation Website. (2018). *Institution classification.* https://carnegieclassifications.iu.edu/downloads/CCIHE2018-FactsFigures.pdf

Chickering, A.W. (1974). *Commuting vs. residential students: Overcoming educational inequities of living off campus.* Jossey-Bass.

Clark, M. R. (2006). Succeeding in the City: Challenges and best practices on urban commuter campuses. *About Campus*, *11*(3), 2–8. https://doi.org/10.1002/abc.166

Council of Economic Advisers. (2016). *Investing in higher education: Benefits, challenges, and the state of student debt.* Washington, DC: Council of Economic Advisers. https://obamawhitehouse.archives.gov/sites/default/files/page/files/20160718_cea_student_debt.pdf

Crede, M., & Niehorster, S. (2012). Adjustment to college as measured by the Student Adaptation to College Questionnaire: A quantitative review of its structure and relationships with correlates and consequences. *Educational Psychology Review*, *24*(1), 33–165.

Creswell, J. W. (2009). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th ed.). SAGE Publications, Inc.

Cooper, H. (1984). *The integrative research review: A systematic approach.* Sage Publications, Inc.

Cunningham, A. F., & Kienzl, G. S. (2011). *Delinquency: The untold story of student loan borrowing*. Institute for Higher Education Policy (IHEP), Washington, DC. http://www.asa.org/pdfs/corporate/delinquency_the_untold_story.pdf

Delisle, J. (2016). Headache for Hillary's higher education plan. *Evidence Speaks Series, 31.*

Department of Education. (2013). *Federal student aid, debt prevention and management*. http://www2.ed.gov/offices/OSFAP/defaultmanagement/cdr.html

Department of Education. (2018a). *Cohort default rate guide.* https://ifap.ed.gov/DefaultManagement/guide/attachments/CDRMasterFile.pdf

Department of Education. (2018b). *National default rate briefings for FY 2015 official cohort default rates.* https://ifap.ed.gov/eannouncements/092618CDRNationalBriefingsFY15.html

Department of Education. (2019a). *Federal student aid.* https://studentaid.ed.gov/sa/repay-loans/default/get-out

Department of Education. (2019b). *Federal student loans: Basics for students.* https://studentaid.ed.gov/sa/sites/default/files/direct-loan-basics-students.pdf

Department of Education. (2019c). *Default rates.* https://studentaid.ed.gov/sa/about/data-center/student/default

Department of Education. (2019d). *Comparison of FY 2016 official national cohort default rates to prior two official cohort default rate.* https://www2.ed.gov/offices/OSFAP/defaultmanagement/schooltyperates.pdf

Department of Education. (2021). *Federal student loan portfolio.* https://studentaid.gov/data-center/student/portfolio.

Dynarski, S. (1994). Who defaults on student loans? Findings from the National Postsecondary Student Aid Study. *Economics of Education Review, 13*(1)*,* 55–68. https://doi.org/10.1016/0272-7757(94)90023-X

Dynarski, S., & Kreisman, D. (2013). Loans for educational opportunity: Making borrowing work for today's students. *Hamilton Project Discussion Paper*. https://www.hamiltonproject.org/papers/loans_for_educational_opportunity

Federal Reserve Bank of New York. (2021). *Consumer credit outstanding (levels).* https://www.federalreserve.gov/releases/g19/HIST/cc_hist_memo_levels.html

Federal Student Aid. (2018a). *Federal Student Grant Programs*. https://studentaid.ed.gov/sa/sites/default/files/federal-grant-programs.pdf

Federal Student Aid. (2018b). *Federal Student Grant Programs: An Office of the U.S. Department of Education*. https://studentaid.ed.gov/sa/glossary#Independent_Student

Flint, T. A. (1997). Predicting student loan defaults. *Journal of Higher Education, 68*(3), 322–354. https://doi.org/10.2307/2960044

Gianoutsos, D. (2011). *Comparing the student profile characteristics between traditional residential and commuter students at a public, research-intensive, urban commuter university* (Doctoral dissertation, University of Nevada, Las Vegas). http://dx.doi.org/10.34917/2264284

Gonzalez Canche, M. S. (2014). Is the community college a less expensive path towards a bachelor's degree? Public 2-year and 4-year colleges' impact on loan debt. *The Journal of Higher Education, 85(5),* 723–759. https://doi.org/10.1080/00221546.2014.11777346

Gross, J. P., Cekic, O., Hossler, D., & Hillman, N. (2009). What matters in student loan default: A review of the research literature. *Journal of Student Financial Aid, 39*(1), 19–29.

Gross, J. P., Hossler, D., Ziskin, M., & Berry, S. B. (2015). Institutional merit-based aid and student departure: A longitudinal analysis. *The Review of Higher Education, 38*(2), 221–250*.*

Harrast, S. A. (2004). Undergraduate borrowing: A study of debtor students and their ability to retire undergraduate loans. *Journal of Student Financial Aid, 34*(1), 21–37.

Heller, D. E. (2008). The impact of student loans on college access. In S. Baum, M. McPherson, & P. Steele (Eds.), *The effectiveness of student aid policies: What the research tells us* (39–68). The College Board.

Heller, D. E. (2011). The financial aid picture: Realism, surrealism, or cubism? In J. C. Smart & M. B. Paulsen (Eds.), *Higher education: Handbook of theory and research* (pp. 125–160). Springer. https://doi.org/ 10.1007/978-94-007-0702-3

Heller, D. E., & Callender, C. (2013). *Student financing of higher education: A comparative perspective*. Routledge.

Herr, E., & Burt, L. (2005). Predicting student loan default for the University of Texas at Austin. *Journal of Student Financial Aid*, *35*(2), 27–49. https://ir.library.louisville.edu/jsfa/vol35/iss2/2/

Hillman N. W. (2014). College on credit: A multilevel analysis of student loan default. *Review of Higher Education*, *37*(2), 169-195. https://doi.org/10.1353/rhe.2014.0011

Hillman, N. W. (2015). Borrowing and repaying student loans. *Journal of Student Financial Aid*, *45*(3), 35–48. http://publications.nasfaa.org/jsfa/vol45/iss3/5

Horn, L., & Nevill, S. (2006). *Profile of undergraduates in U.S. postsecondary education institutions: 2003–04: With a special analysis of community colleges students.* National Center for Education Statistics. http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2006184

Hylands, T. (2014). *Student loan trends in the third federal reserve district.* Cascade Focus.

Hu, S., & St. John, E. P. (2001). Student persistence in a public higher education system. *The Journal of Higher Education*, *72*(3), 265–286. https://doi.org/10.1080/00221546.2001.11777095

Ishitani, T. T., & Reid, A. M. (2015). First-to-Second-Year Persistence Profile of Commuter Students. In J. P. Biddix (Ed.), *Understanding and addressing commuter student needs: New directions for student services*, *number 150*, (pp. 13–26). Wiley Periodicals. https://doi.org/10.1002/ss.20123

Jabareen, Y. (2009). Building a conceptual framework: Philosophy, definitions, and procedure. *International Journal of Qualitative Methods, 8*(4), 49-62. https://doi.org/10.1177/160940690900800406

Jacoby, B. (2000). Why involve commuter students in learning? In M. Kramer (Series Ed.), & B. Jacoby (Vol. Ed.). *New directions for higher education: Number 109. Involving commuter students in learning* (pp. 3-12). Jossey-Bass.

Jacoby, B. (2015). Enhancing commuter student success: What's theory got to do with it? *New Directions for Student Services*, *150*, 3-12. https://doi.org/ 10.1002/ss.20122

Jacoby, B., & Garland, (2004). Strategies for enhancing commuter student success. *Journal of College Student Retention: Research, Theory and Practice*, *6*(1), 61–79.

Johnson, C. L. (2012). *Do new student loan borrowers know what they are signing? A phenomenological study of the financial aid experiences of high school seniors and college freshman* (Doctoral dissertation, Iowa State University). https://lib.dr.iastate.edu/cgi/viewcontent.cgi?referer=https://www.google.com/&httpsredir=1&article=3362&context=etd

Johnson, J. L. (1997). Commuter college students: What factors determine who will persist and who will drop out? *College Student Journal, 31*(3), 323–332.

Kane, T. J. (1999). *The price of admission.* Russell Sage Foundation.

Keeling, R. P. (1999). A new definition of college emerges: Everything that happens to…a (newly defined) student, in the context of a noisy visual 'datascape.' *NASPA Forum*, *20* (5), 4–5.

Knapp, L. G., & Seaks, T. G. (1992). An analysis of the probability of default on federally guaranteed student loans. *The Review of Economics and Statistics, 74*(3), 404–411.

Kuh, G. D., Gonyea, R. M., Palmer, M. (2001). The disengaged commuter student: Fact or fiction? *Commuter Perspectives, 27*. https://hdl.handle.net/2022/24256

Kuh, G. D., Kinzie, J., Buckley, J. A., Bridges, B. K., & Hayek, J. C. (2007). *Piecing together the student success puzzle: Research, propositions, and recommendations* [ASHE-ERIC Higher Education Report, Vol. 32, No. 5]. Jossey-Bass.

Kuzma, A. T., Kuzma, J. R., & Thiewes, H. F. (2010). An examination of business students' student loan debt and total debt. *American Journal of Business Education, 3*(4), 71–78. https://doi.org/10.19030/ajbe.v3i4.416

Lomax, R. G., & Hahs-Vaughn, D. L. (2012). *Statistical concepts* (3rd ed.). Routledge.

Looney, A. & Yannelis, C. (2015). A crisis in student loans? How changes in the characteristics of borrowers and in the institutions they attended contributed to rising loan defaults. *Brookings Papers on Economic Activity*. https://www.brookings.edu/bpea-articles/a-crisis-in-student-loans-how-changes-in-the-characteristics-of-borrowers-and-in-the-institutions-they-attended-contributed-to-rising-loan-defaults/

Lund Research Ltd. (2012). *Total population sampling.* Laerd Dissertation. https://dissertation.laerd.com/total-population-sampling.php#adv-dis

Marshall, C., & Rossman, G. B. (2006). *Designing qualitative research* (4th ed.). Sage.

Mayhew, M. J., Rockenback, A. N., Bowman, N. A., Seifert, T. A., Wolniak, G. C., Pascarella, E. T., & Terenzini, P.T. (2016). *How college affects students: Volume 3: 21st century evidence that higher education works.* Jossey-Bass.

Melendez, M. C. (2016). Adjustment to college in an urban commuter setting: The impact of gender, race/ethnicity, and athletic participation. *Journal of College Student Retention: Research, Theory and Practice*, *18*(1), 31–48. https://doi.org/10.1177/1521025115579671

Mezza, A. A., & Sommer, K. (2015). A trillion dollar question: What predicts student loan delinquencies? *Finance and Economics Discussion Series, 2015-098.* Washington: Board of Governors of the Federal Reserve System. http://dx.doi.org/10.17016/FEDS.2015.098

Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded source book* (2nd ed.). Sage.

Monks, J. (2012). The role of tuition, financial aid policies, and student outcomes on average student debt (Working paper). University of Richmond. https://www.ilr.cornell.edu/cheri/workingPapers/upload/cheri_wp150.pdf

Monteverde, K. (2000). Managing student loan default risk: Evidence from a privately guaranteed portfolio. *Research in Higher Education, 41*(3), 331–352.

Naes, T., Tomic, O., Afseth, N., Segtnan, V., & Mage, I. (2013) Multi-block regression based on combinations of orthogonalisation, PLS-regression and canonical correlation analysis. *Chemometrics and Intelligent Laboratory Systems, 124*, 32–42. 10.1016/j.chemolab.2013.03.006

National Association of Student Financial Aid Administrators (NASFAA). (2019). *Data sharing decision tree updated to reflect expanded allowable data sharing*. https://www.nasfaa.org/news-item/17144/Data_Sharing_Decision_Tree_Updated_to_Reflect_Expanded_A

National Association of Student Financial Aid Administrators (NASFAA). (2021). *National student loan cohort default rate continues to drop*. https://www.nasfaa.org/news-item/23473/National_Student_Loan_Cohort_Default_Rate_Continues_to_Drop

National Center for Education Statistics (NCES). (2013). *The condition of education: 2013*. http://nces.ed.gov/programs/coe/indicator_cud.asp

National Center for Education Statistics (NCES). (2014). *IPEDS Analytics: Delta Cost Project Database.* http://nces.ed.gov/ipeds/deltacostproject/

National Center for Educational Statistics (NCES). (2017). *The condition of education*. https://nces.ed.gov/fastfacts/display.asp?id=31

National Center for Educational Statistics (NCES). (2018a). *The condition of education*. https://nces.ed.gov/fastfacts/display.asp?id=98

National Center for Educational Statistics (NCES). (2018b). *Digest of education statistics*. https://nces.ed.gov/programs/digest/d16/tables/dt16_303.70.asp?current=yes

National Center for Educational Statistics (NCES). (2018c). *Digest of education statistics*. https://nces.ed.gov/programs/digest/d16/tables/dt16_303.80.asp?current=yes

National Center for Educational Statistics (NCES). (2019). *Integrated Postsecondary Education Data System (IPEDS) Data Center*. U.S. Department of Education, Institute of Education Sciences. https://nces.ed.gov/ipeds/datacenter/

National Center for Educational Statistics (NCES). (2019). *Integrated Postsecondary Education Data System (IPEDS) Data Center*. U.S. Department of Education, Institute of Education Sciences. https://nces.ed.gov/pubs/web/97578e.asp

National Center for Educational Statistics (NCES). (2020). *Integrated Postsecondary Education Data System (IPEDS) Data Center*. U.S. Department of Education, Institute of Education Sciences.

National Student Loan Data System (NSLDS). (2020). *Retrieve your loan information*. https://nslds.ed.gov/nslds/nslds_SA/

National Survey of Student Engagement (NSSE). (2017). *Evidence based improvement in higher education.* https://nsse.indiana.edu/html/summary_tables.cfm

Pallant, J. (2010). *SPSS survival manual* (5th ed.). Open University Press.

Pascarella, E. T., & Terenzini, P.T. (2005). *How college affects students (Volume 2): A third decade of research.* Jossey-Bass.

Paulsen, M. B., & St. John, E. P. (2002). Social class and college costs: Examining the financial nexus between college choice and persistence. *Journal of Higher Education*, *73*(2), 189–236. https://doi.org/10.1080/00221546.2002.11777141

Podgursky, M., Ehlert, M., Monroe, R., Watson, D., & Wittstruck, J. (2002). Student loan defaults and enrollment persistence. *Journal of Student Financial Aid, 32*(3), 27–42.

Rothstein, J., & Rouse, C. E. (2011). Constrained after college: Student loans and early-career occupational choices. *Journal of Public Economics*, *95*(1–2), 149–163. https://doi.org/10.1016/j.jpubeco.2010.09.015

Rhyneer, M. (2021, September 14). *3 test-optional takeaways for enrollment leaders from the 2021 admissions cycle*. EAB. https://eab.com/insights/blogs/enrollment/test-optional-takeaways-for-enrollment-leaders/

Scott-Clayton, J. (2018). The looming student loan default crisis is worse than we thought. *Evidence Speaks Reports, 2*(34).

Schlossberg, N. K., Lynch, A. Q., & Chickering, A.W. (1989). *Improving higher education environments for adults*. Jossey-Bass.

Steiner, M., & Teszler, N. (2003). *The characteristics associated with student loan default at Texas A&M University.* Texas Guaranteed Student Loan Corporation.

Steiner, M., & Teszler, N. (2005). *Multivariate analysis of student loan defaulters at Texas A&M University.* Texas Guaranteed Student Loan Corporation.

Tinto, V. (1975). Dropout from higher education: A theoretical synthesis of resent research. *Review of Educational Research*, *45*(1), 89–125. https://doi.org/10.3102/00346543045001089

Tinto, V. (1993). *Leaving college: Rethinking the causes and cures of student attrition* (2nd ed.). The University of Chicago Press.

Tinto, V. (2006). Research and practice of student retention: What next? *Journal of College Student Retention: Theory, Research, and Practice*, *8*(1), 1–19. https://doi.org/10.2190/4YNU-4TMB-22DJ-AN4W

Torres, V. (2006). A mixed method study testing data-model fit of a retention model for Latino/a students at urban universities. *Journal of College Student Development*, *47*(3), 299–318. https://doi.org/10.1353/csd.2006.0037

Volkwein, J. F., & Szelest, B. P. (1995) Individual and campus characteristics associated with student loan default. *Research in Higher Education, 36*(1), 41–72. https://doi.org/ 10.1007/BF02207766

Volkwein, J. F., Szelest, B. P., Cabrera, A. F., & Napierski-Prancl, M. R. (1998). Factors associated with student loan default among different racial and ethnic groups. *Journal of Higher Education, 69*(2), 206–237. https://doi.org/10.1080/00221546.1998.11775133

Wilms, W. W., Moore, R. W., & Bolus, R. E. (1987). Whose fault is default? A study of the impact of student characteristics and institutional practices on guaranteed student loan default rates in California. *Educational Evaluation and Policy Analysis, 9*(1), 41–54. https://doi.org/ 10.3102/01623737009001041

Wilmes, M. B., & Quade, S. L. (1986). Perspectives on programming for commuters: Examples of good practice. *NASPA Journal*, *24*(1), 25–35. https://doi.org/10.1080/00220973.1986.11071983

Woo, J. H. (2002). Factors affecting the probability of default: Student loans in California. *Journal of Student Financial Aid, 32*(2), 5–23. https://ir.library.louisville.edu/jsfa/vol32/iss2/1